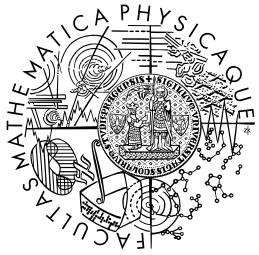

FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

MASTER THESIS

Mikuláš Matoušek

Electronic structure calculations of biologically relevant transition metal complexes

Department of chemical physics and optics

Supervisor of the master thesis: RNDr. Libor Veis, Ph.D.
Study programme: Physics
Study branch: Biophysics & Chemical Physics

Prague 2020

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In ..... date ....

Author's signature

Given by its complexity, this work could never have been finished by myself alone. I therefore owe a huge thank you to many people. Not surprisingly, the most important man in the process was my supervisor, Libor Veis, who not only was there for me providing helpful suggestions whenever I needed, but also created an excellent work environment for me. And most importantly, without his enthusiasm providing me with much needed motivation I would have long lost the will to finish this work.

I have also received a lot of help from Andrej Antalík, not only with technical issues surrounding all the used software, as he is very skilled in these areas and always eager to help, while, conveniently, sharing an office with me, but he also has a significant contribution to the part regarding the porphyrin model calculations.

I am glad that Martin Srnec invited me to collaborate on his project regarding IPCs. I would like to thank Kasia Pernal and Michal Hapka of the Łódź University of Technology, who not only introduced me to the adiabatic connection method and provided me with their code, but also invited me for a lovely visit in Poland in order to explain how to use it. I would like to thank Yang Guo for helping with the AC0 calculations.

A minor, but crucial contribution was the one of Honza Brandejs, who brought me to the J. Heyrovský Institute, which I am extremely grateful for.

Many others, while not contributing directly, have provided support for me, which was just as necessary to finish this work.

Title: Electronic structure calculations of biologically relevant transition metal complexes

Author: Mikuláš Matoušek

Department: Department of chemical physics and optics

Supervisor: RNDr. Libor Veis, Ph.D., J. Heyrovský Institute of Physical Chemistry of the CAS, v. v. i.

Abstract: Porphyrins are an important class of biomolecules, which are heavily studied, both experimentally and computationally. But, despite the intensive efforts, for many questions we still aren't able to consistently find an agreement between theory and experiment. One of the still unresolved issues is the character of the ground state of the Fe(II)-porphyrin molecule. We used a model of the Fe(II)-porphyrin molecule to study the effects of geometrical changes on the spin states. By carrying out extensive DMRG-CASSCF calculations topped with TCCSD correlation treatment we are able to link the effects of these geometrical changes to the experimental results, and predict a quintet ground state for the isolated Fe(II)-porphyrin molecule. Also, using a ligated porphyrin belonging to the iron porphyrin carbene class of molecules, we demonstrate by combining the CASSCF and AC0 methods that geometrical changes outside the porphyrin core cannot be overlooked.

Keywords: Fe(II) porphyrins, Quantum chemical calculations, DMRG, Coupled cluster methods, Adiabatic connection

Contents

List of Abbreviations

AC	–	Adiabatic Connection
CAS	–	Complete Active Space
CASSCF	–	Complete Active Space Self Consistent Field
CCSD	–	Coupled Clusters Singles and Doubles
DBSS	–	Dynamic Block State Selection
DFT	–	Density Functional Theory
DLPNO-TCCSD	–	Domain based Local Pair Natural Orbital TCCSD
DMRG	–	Density Matrix Renormalization Group
ERPA	–	Extended Random Phase Approximation
FCI	–	Full Configuration Interaction
HF	–	Hartree-Fock
IPC	–	Iron Porphyrin Carbene
MCSCF	–	Multi Configurational SCF
MPS	–	Matrix Product State
PNO	–	Pair Natural Orbital
RDM	–	Reduced Density Matrix
SA	–	State Averaged
SCF	–	Self Consistent Field
SD	–	Slater Determinant
SS	–	State Specific
SVD	–	Singular Value Decomposition
TCC	–	Tailored Coupled Clusters
TCCSD	–	Tailored Coupled Clusters Singles and Doubles
TDM	–	Transition Denstity Matrix

Introduction

Porphyrins and porphyrin-like molecules are a broad class of compounds of huge interest. They consist of 4 fused tetrapyrrole rings linked by methylidyne bridges forming a larger macrocycle. The molecule has a conjugated π-system, spreading over the whole molecule.

They bear a crucial role in biochemistry, mostly as cofactors of proteins in many important biochemical processes, including molecular transport, enzymatic catalysis and photosynthesis [1]. This versatility is given by their unique ability to exist in different oxidation and spin states, having many electronic states very close to each other and the ability to fine tune their properties by changing the encapsulating protein or the surrounding environment. The most well known example is probably the heme molecule in hemoglobin, which is responsible for the distribution of oxygen over the body of many higher organisms.

This flexibility not only leads to their widespread occurrence in most organisms, but also enables many applications in engineering. They are proposed for use in materials for photonics, as a part of liquid crystals, in Metal Organic Frameworks, conductive polymers or as a photosentitizer in photodynamic therapy [2, 3]. They are also investigated as cofactors in chemical synthesis [4]. Their potential use is further extended by the fact that while living organisms mostly make use of only Fe-porphyrins in synthetic applications we can use the porphyrin skeleton to accommodate cations of many transition metals throughout the periodic table.

This class of molecules is challenging to study computationally and is problematic for the direct comparison with experiment. The issues on the side of experiment arise from the instability of the Fe(II)-porphyrin molecule, which exists only when stabilized by sidechains on the porphyrin circle and/or axial ligands and in biochemistry by the surrounding protein. On the other side, they can be studied using a wide range of experimental methods. They have strong absorption bands in the UV/visible regions, called the Soret bands. Most of the

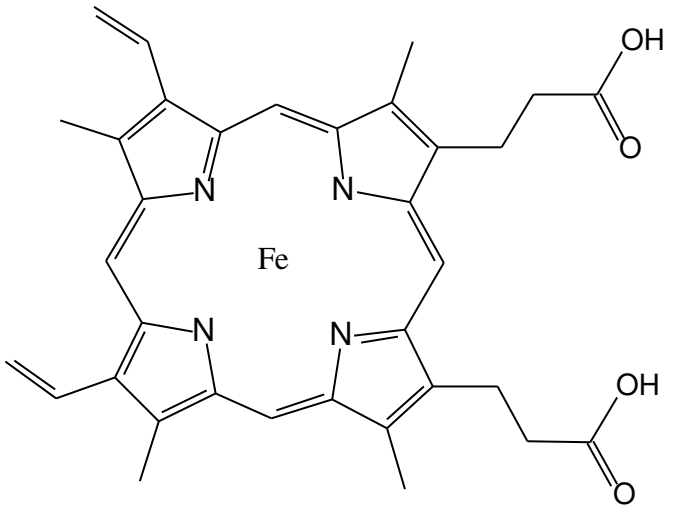

Figure 1: The heme B molecule, being the carrier of oxygen, is probably the most well known example of a porphyrin system.

biologically most relevant porphyrin systems contain an iron atom, which enables them to be studied using Mössbauer spectroscopy. Other options include X-ray diffractometry, EPR NMR or Raman spectroscopies and many others [5, 6, 7, 8].

Electronic structure calculations typically fail on the fact, that the molecule is strongly correlated, with the active space necessary to describe the correlation significantly larger than what is possible by current conventional CASSCF [9, 10]. What makes calculations even more challenging is that for the Fe(II)-porphyrin molecule, the high-spin, intermediate-spin and low-spin states of the molecule are energetically very close to each other, which leads to unsureness about the correct spin state ordering, which is dependent on the used computational method [9, 11].

A good illustration of the struggles would be the spin crossover in the heme system, which, upon the binding of oxygen, has to rapidly undergo a spin forbidden transition, the mechanism behind which is not yet fully known, despite being intensively studied [12].

This motivated us to do a study on these systems, as not only the compounds are of great interest, but more importantly, the multireference nature of the porphyrin core [13] serves as an excellent problem where we can make perfect use of the multireference computational methods developed in our group.

Also, as many of these systems have been heavily studied computationally, there is a plenty of results obtained using various quantum chemistry methods, which can serve as the much needed reference data to assess the accuracy and reliability of the methods we have developed.

We have decided to study two porphyrin systems, a simplified model of the porphyrin core, well suited for accurate calculations, and a larger system more close to potential appplications. Both of these systems are decribed in much more details later in this work.

1. Theoretical background

Throughout this work, we will use natural units which are defined so that

$$\hbar = m_e = \frac{1}{4\pi\epsilon_0} = 1.$$

Quantum chemistry is interested in solving the Schrödinger equation for molecular systems. As this is generally impossible to do analytically and difficult to perform numerically, it is necessary to employ a series of approximations.

The first usually used are the Born-Oppenheimer and adiabatic approximations, which, due to a large difference in the masses of electrons and nuclei, uncouple the nuclear and electronic motion, enabling the problems to be solved separately and nuclear coordinates to be kept constant throughout the solution of the electronic problem.

This eventually leads to the electronic Hamiltonian

$$\hat{H} = -\sum_i \frac{1}{2}\Delta_i - \sum_{I,i} \frac{Z_I}{|\vec{R}_I - \vec{r}_i|} + \sum_{I,J} \frac{Z_J Z_I}{|\vec{R}_I - \vec{R}_J|} + \sum_{i,j:i<j} \frac{1}{|\vec{r}_i - \vec{r}_j|}, \tag{1.1}$$

where the capital indices relate to the the nuclear coordinates $\vec{R}$ and nuclear charges Z, while the lowercase indices are related to the electronic coordinates $\vec{r}$.

1.1 The Hartree-Fock approximation

Solving the Schrödinger equation for the electronic Hamiltonian still involves finding a solution of a $3N$-dimensional partial differential equation for the N-electron wave function, which usually is not possible, except for very small Ns (in the order of units).

The solution to this problem is to consider the solution as a fully antisymmetrized product of single electron functions called the spin-orbitals, which is written in the form of a Slater determinant

$$\psi_{\mathrm{SD}}(\vec{x_1}, \vec{x_2}, \cdots, \vec{x_N}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\vec{x}_1) & \chi_2(\vec{x}_1) & \cdots & \chi_N(\vec{x}_1) \\ \chi_1(\vec{x}_2) & \chi_2(\vec{x}_2) & & \chi_N(\vec{x}_2) \\ \vdots & & \ddots & \vdots \\ \chi_1(\vec{x}_N) & \chi_2(\vec{x}_N) & \cdots & \chi_N(\vec{x}_N) \end{vmatrix}. \tag{1.2}$$

Here $\vec{x}$ denotes the combination of the three spatial and the spin variable.

Using this ansatz in the time independent Schrödinger equation eventually leads to the Hartree-Fock equations.

$$\hat{F}\chi_i(\vec{x}) = \epsilon_i \chi_i(\vec{x}), \tag{1.3}$$

With the explicit form of the Fock operator acting on the spin orbital being

$$\begin{aligned} \hat{F}\chi_i(\vec{x}) = \frac{1}{2}\Delta\chi_i(\vec{x}) - \sum_I \frac{Z_I}{|\vec{R}_I - \vec{r}|}\chi_i(\vec{x}) + \sum_{j,j\neq i} \int \frac{\chi_j^*(\vec{x}')\chi_j(\vec{x}')}{|\vec{r} - \vec{r}'|}\chi_i(\vec{x})\,\mathrm{d}\vec{x}' - \\ - \int \frac{\chi_j^*(\vec{x}')\chi_i(\vec{x}')}{|\vec{r} - \vec{r}'|}\chi_j(\vec{x})\,\mathrm{d}\vec{x}'. \end{aligned} \tag{1.4}$$

But, due to electron-electron repulsion, the explicit form of the Fock operator $\hat{F}$ is dependent on the spinorbitals $\chi_i(\vec{x})$, making the resulting equations nonlinear. Therefore we need to solve them iteratively in a self-consistent manner. The numerical solving of these equations can be further simplified by separating the spatial and spin parts of the spinorbitals and integrating out the spin part, leaving us with the equations for the spatial part $\psi_i(\vec{r})$ of the wave function, which then is further simplified by introducing an atomic basis set

$$\psi_i(\vec{r}) = \sum_\mu \mathbb{C}_{i\mu}\phi_\mu(\vec{r}), \tag{1.5}$$

for a fixed set of basis functions ϕ_μ. This, for the most often encountered closed shell singlet state, leads to the Roothan equations, which take a matrix form

$$\mathbb{FC} = \mathbb{SC}\mathcal{E}, \tag{1.6}$$

where the matrix $\mathbb{F}$ is the Fock operator represented in the atomic orbital basis

$$(\mathbb{F})_{\mu\nu} = \int \phi_\mu^*(\vec{r})\hat{F}\phi_\nu(\vec{r})\,\mathrm{d}\vec{r}, \tag{1.7}$$

and $\mathbb{S}$ is the basis set overlap matrix

$$(\mathbb{S})_{\mu\nu} = \int \phi_\mu^*(\vec{r})\phi_\nu(\vec{r})\,\mathrm{d}\vec{r}. \tag{1.8}$$

The solutions of these equations are the orbital energies $\mathcal{E}$ in the form of a matrix, diagonal by convention, and the matrix of expansion coefficients $\mathbb{C}$, representing the HF orbitals expanded in the basis set. Due to the double counting of the electron repulsion term in the single orbital energies, the HF energy of the full system is

$$E_{\mathrm{HF}} = \sum_i \epsilon_i - \frac{1}{2}\sum_{ij} \langle ij|ij\rangle - \langle ij|ji\rangle\,, \tag{1.9}$$

where we introduce the commonly used Dirac notation of two electron integrals.

$$\langle ij|kl\rangle = \int \chi_i^*(\vec{x_1})\chi_k(\vec{x_1})\frac{1}{|\vec{r}_1 - \vec{r}_2|}\chi_j^*(\vec{x_2})\chi_l(\vec{x_2})\,\mathrm{d}\vec{x_1}\,\mathrm{d}\vec{x_2} \tag{1.10}$$

The Hartree-Fock method is in much more detail explained in [14].

1.1.1 Basis sets

Over the years, quantum chemistry has settled on using Gaussian functions for the basis set. Each basis set function is either a Gaussian with a fixed exponent, or a linear combination of several such functions. While this choice does not guarantee the optimal convergence with respect to the basis set size, they offer a massive speedup due to being able to find an analytical form of the one and two electron integrals used to construct the Fock operator matrix.

1.1.2 Correlation energy

While the Hartree-Fock approximation can be pretty reasonable for a broad class of molecules, the error introduced to the energy is $\sim 1\%$. The inaccuracy is mostly due to the inability to describe the correlation of positions of the different electrons with opposite spins by the single Slater determinant.

This leads to the introduction of the concept of correlation energy, which is defined as the difference between the exact (nonrelativistic) energy, and the energy from the Hartree-Fock method

$$E_{\mathrm{Corr}} = E_{\mathrm{Exact}} - E_{\mathrm{HF}}, \tag{1.11}$$

in the limit of a complete basis set.

To account for the correlation energy, we have a wide range of options. Apart from the widely used density functional theory (DFT), most other schemes to obtain the correlation energy rely in some way on the expansion of the solution in term of Slater determinants. As these methods usually use determinants constructed from the Hartree-Fock orbitals, they are called post-Hartree-Fock methods.

Usually, we distinguish two types of electronic correlation. Dynamic correlation, where the correlation energy comes from small contributions of many Slater determinants, and static correlation, where there are several determinants which have strong contributions.

While in the first case the Hartree-Fock method serves as a very good starting point for post-HF calculations, static correlation implies a small expansion coefficient of the HF determinant in the exact wave function which is consequently a poor reference for the most often used schemes to obtain the correlation energy, and strongly correlated systems generally call for the use of the so-called multireference methods.

1.1.3 FCI expansion

The exact solution can be found by expanding the wave function in the basis of the full Hilbert space. This is best done in the form of Slater determinants. Denoting $|\psi_0\rangle$ the Hartree-Fock determinant, $|\psi_i^a\rangle$ the determinant obtained by exciting the electron from the i-th to the a-th spinorbital in the HF determinant, and similarly $\left|\psi_{ij}^{ab}\right\rangle$,$\left|\psi_{ijk}^{abc}\right\rangle$ etc. the multiply excited determinants, we can write the wave function as

$$|\Psi\rangle = c_0 |\psi_0\rangle + \sum_{i,a} c_i |\psi_i^a\rangle + \sum_{ij,ab} c_{ij} \left|\psi_{ij}^{ab}\right\rangle + \ldots, \tag{1.12}$$

and then solve the Schrödinger equation to find the c coefficients. Unfortunately, the dimensionality of the space grows exponentially with respect to the number of electrons rendering the method, called full configuration interaction (FCI), prohibitively expensive for most problems. This leads to many schemes to try to approximate the FCI result.

1.1.4 Density matrices

It is often useful to define the density operator. For a pure quantum state defined $|\Psi\rangle$, the density operator is defined by

$$\hat{\rho} = |\Psi\rangle \langle\Psi| , \tag{1.13}$$

which offers a full description of the state of the system. In quantum chemistry, it is common to use the positional representation by some wave function $\Psi(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N)$. In this representation, the density operator takes on the form

$$\hat{\rho}(\vec{x}_1, \vec{x_1}', \vec{x}_2, \vec{x}_2', \ldots, \vec{x}_N, \vec{x}_N') = \Psi(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N)\Psi(\vec{x}_1', \vec{x}_2', \ldots, \vec{x}_N'). \tag{1.14}$$

The reduced density operators are formed by integrating out the position and spin variables of some of the electrons. The one electron reduced density operator is defined by

$$\rho_1(\vec{x}_1, \vec{x}_1') = N \int \rho(\vec{x}_1, \vec{x}_1', \vec{x}_2, \vec{x}_2, \ldots, \vec{x}_N, \vec{x}_N)\, \mathrm{d}\vec{x}_2 \ldots \mathrm{d}\vec{x}_N \tag{1.15}$$

Similarly, with different prefactors and less integration out variables are defined the 2, 3... electron reduced density operators. Representing the density operators using a basis set results in a matrix form, called the density matrix and the reduced density matrices.

The diagonalization of the one electron reduced density matrix gives us the natural occupation numbers as the eigenvalues and the natural orbitals (precisely their basis set expansion coefficients) as the eigenstates.

1.2 The Density matrix renormalization group method

One of the methods approximating the FCI method is the DMRG method, which emerged quite recently. [15]. Although the method was introduced a little differently, it essentially uses the matrix product state (MPS) ansatz [16], which is used as an approximation of the FCI coefficients, that can be understood as a tensor of high dimensionality. This can be shown by switching to the occupation representation of the Slater determinants, given by $|n_1, n_2, \ldots, n_k\rangle$, where n_i denotes the occupation of the i-th orbital and can acquire 4 values, as the orbital can be either doubly occupied, empty, or singly occupied with either of the two spins. The FCI wave function then has the form

$$|\Psi\rangle = \sum_{\{n\}} c^{n_1, n_2, \ldots, n_k} |n_1, n_2, \ldots, n_k\rangle , \tag{1.16}$$

with the occupancies $\{n\}$ fulfilling the condition

$$\sum_i |n_i| = N.$$

The MPS ansatz approximates this tensor $c^{n_1, n_2, \ldots, n_k}$ as a product of rank-3 tensors, with one index being called physical and corresponding to one of the

indices of the original FCI tensor, while the other two, called virtual indices, are contracted with other tensors in the MPS.

$$|\Psi\rangle = \sum_{\{n\},i,j,\ldots,q,r} A_i^{n_1} A_{ij}^{n_2} \ldots A_{qr}^{n_{k-1}} A_r^{n_k} \, |n_1, n_2, \ldots, n_k\rangle \, . \tag{1.17}$$

The contractions of the neighboring tensors give rise to correlations between orbital occupancies. The FCI tensor and its MPS approximation, using the Penrose notation [17, 18], are shown in figure 1.1.

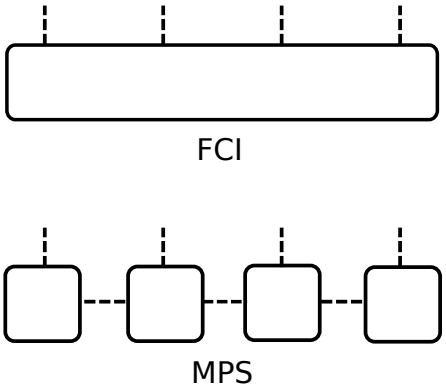

Figure 1.1: The FCI tensor and the MPS ansatz, written using the Penrose notation. In this notation, tensors are represented by squares or rectangles and the outgoing (dashed) lines, often called legs, represent indices. If two tensors are connected by a line, it represents a contraction over those indices.

But the exponential scaling of the original FCI tensor would still manifest itself in the exponential growth of the necessary dimension of the contracted virtual indices. As this would lead to scaling equal that of the original FCI method, we have to limit the maximum dimension of the virtual indices, also called the bond dimension. This reduces the number of variational parameters from $\mathcal{O}(4^k)$ of the original FCI tensor to $\mathcal{O}(4kM^2)$. This is achieved by means of singular value decomposition (SVD) of the tensors in the MPS and keeping only a number of the largest singular values, according to the required bond dimension. The sum of the discarded eigenvalues is then called the truncation error and is related to the error in energy.

The maximum bond dimension is a parameter of the method, which determines both its accuracy and computational cost. Another commonly used approach would be to exploit the fact that the dimension of the virtual indices does not have to be the fixed over the MPS and instead specify a value of the truncation error to be kept constant over the MPS. This is called the Dynamic Block State Selection approach (DBSS) [19].

The structure of the MPS ansatz can be further exploited by writing the Hamiltonian in the second quantization formalism

$$\hat{H} = \sum_{ij} h_{ij} \hat{a}_i^\dagger \hat{a}_j + \frac{1}{2} \sum_{ijkl} \langle ij|kl\rangle \, \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k \tag{1.18}$$

in a tensor network form, also called a Matrix Product Operator (MPO) [16]. This form allows us, by evaluating the contractions of indices in a convenient

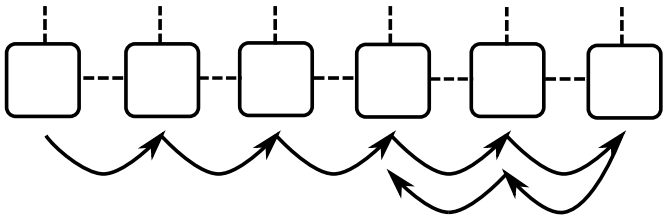

Figure 1.2: An illustration of the sweeping procedure. Note that for the two-site DMRG we optimize two neighboring sites at a time.

order, to efficiently calculate the expectation values of operators, especially the energy value $\langle\Psi|\,\hat{H}\,|\Psi\rangle$.

The solution of the eigenvalue problem itself is done in an iterative procedure, where the matrices in the MPS are optimized. In the DMRG procedure, we optimize one or two matrices of the MPS at a time. The two different approaches are called one-site and two-site DMRG. The optimization goes back and forth over the tensors in the MPS, and is called sweeping. While the single-site DMRG is variational and exhibits strictly monotonic convergence, both of which hold only approximately for the two-site algorithm, it is heavily prone to getting stuck in local minima compared to its two-site counterpart. This renders the two-site approach much more practical.

The whole process proceeds in the following manner:

1. The Hamiltonian is applied to the MPS wave function.

2. The tensor network corresponding to $\langle\Psi|\,\hat{H}\,|\Psi\rangle$ is contracted, except for the one or two sites to be optimized, forming an effective Hamiltonian for the site(s), called the superblock Hamiltonian. (See Figure 1.3) For the two-site algorithm, the two tensors are contracted together and treated as a single tensor.

3. The eigenstates for the superblock Hamiltonian are found, usually by the iterative Davidson algorithm [20], and are used to replace the old tensors at the studied sites.

4. For the two site algorithm, the resulting two-site tensor is split by SVD back into two single site tensors.

5. We move one site to the left/right and repeat the whole procedure

The procedure is considered to be converged when the energy change between sweeps falls below a set threshold.

The computational scaling of the DMRG method is $O(M^3n^3)+O(M^2n^4)$, with the first term usually being dominant [21]. Here M denotes the bond dimension, while n is the number of orbitals in the active space. The method itself is both variational and size-consistent. (Although the latter only holds if the ordering of the MPS tensors is consistent with the division into subsystems.) As the HF determinant is not privileged in any way, the method is considered to be truly multireference, without any fixed reference determinant.

While the 1D nature of the MPS parametrization causes the method to be best suited for 1-D systems, it is routinely used for general molecular systems, at the

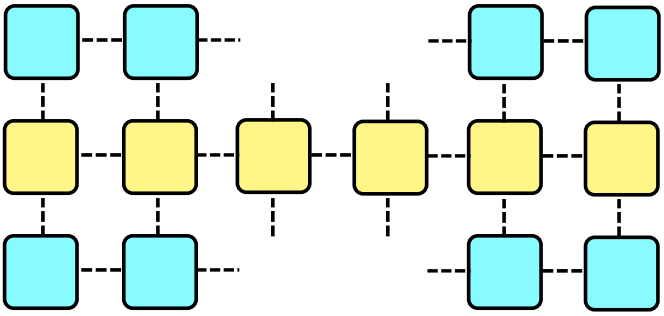

Figure 1.3: The superblock Hamiltonian written in Penrose notation. This is the form for the two site algorithm. In yellow is the two MPO Hamiltonian representation, while in blue are the parts of the MPS form of the bra and ket wavefunctions, not belonging to the current active site.

cost of increasing the required bond dimension to preserve accuracy. Currently, for molecules the method allows calculations containing up to around 50 orbitals in the active space [22], with the precise number depending on the molecular topology, amount of correlation in the molecule, and the required accuracy.

1.2.1 Mutual information

In many body physics, the von Neumann entropy of a state described by a density matrix ρ is given by

$$S(\rho) = -\mathrm{Tr}\,\rho \log \rho \tag{1.19}$$

In DMRG we use this concept to define the one orbital (site) entropy

$$S_i = \mathrm{Tr}\,\rho_i \log \rho_i, \tag{1.20}$$

$$\rho_i = \mathrm{Tr}_{k \neq i}\,\rho,$$

where the trace in the density matrix is over all orbitals (sites) in the MPS except the one of current interest. This is schematically shown in Figure 1.4. Similarly, the two orbital entropy is defined by

$$S_{ij} = \mathrm{Tr}\,\rho_{ij} \log \rho_{ij}, \tag{1.21}$$

$$\rho_{ij} = \mathrm{Tr}_{k \neq i,j}\rho,$$

where we omit two orbitals in the tracing.

These can be used to find the mutual information of two orbitals, using the definition

$$I_{ij} = S_i + S_j - S_{ij} \tag{1.22}$$

This measure can be related to the correlation between the electrons in the two orbitals and is also used to optimize the ordering, mentioned in the next subsection.

The one orbital entropy is often used to choose the orbitals in the active space used in multireference calculations [23].

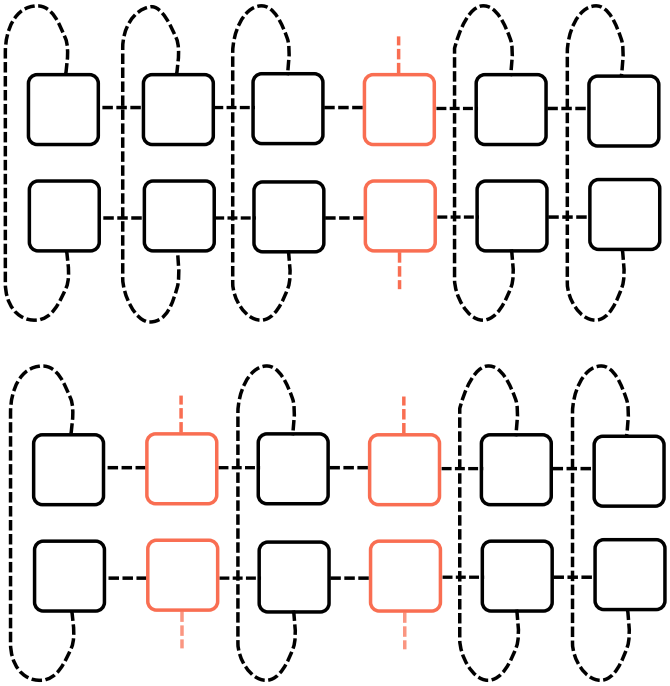

Figure 1.4: An example of the contractions of the MPS used to obtain the ρ_i (top) and ρ_{ij} (bottom), necessary to find the mutual information of orbitals.

1.2.2 Orbital ordering

The DMRG method, if the bond dimension is limited, is not invariant to changing the order of the orbitals in the MPS, as the error introduced by truncating the bond dimensions depends on the total mutual information "going over" the contracted index in the MPS.

This motivates us to try to minimize the introduced error. That is done by choosing the order of orbitals such that the orbitals with high mutual information are closer to each other, which is achieved by minimizing the cost function

$$\mathrm{COST}_\eta = \sum_{ij} I_{ij} d_{ij}^\eta, \tag{1.23}$$

where d_{ij} is the distance between the sites and η is chosen mostly arbitrarily. For $\eta = 2$ we can use the Fiedler algorithm from spectral graph theory [24], or we can use genetic algorithms, which do not impose a restriction on the value of η.

1.3 CASSCF

Many post-HF methods rely on the fact that a single Slater determinant is the dominant contribution to the exact solution. Although this often holds, it is not always the case, causing the methods to breakdown on systems with strong static correlation.

The most common approach to deal with strongly correlated systems is to have a linear combination of Slater determinants as the initial guess and optimize both the orbitals and the coefficients of the expansion. This is generally called Multi-Configurational SCF (MCSCF).

A very commonly used choice for the expansion is to take all possible determinants from a set of orbitals, thus effectively performing a FCI in a subspace (called the active space) of the full Hilbert space, coupled with an orbital optimization step. This is called the Complete Active Space Self Consistent Field method (CASSCF).

The size of the active space is usually demoted CAS(X, Y), where X denotes the number of electrons and Y the number of orbitals in the active space.

For calculating excited states we have several options. Either we can use separate calculations for each of the states, this being called the State Specific (SS) approach, or we perform the whole calculation and orbital optimization using the average electronic density of all states in which we are interested, leading to the State Averaged (SA) approach.

The results obtained by this method are not invariant to the choice of the active space. In order to obtain reliable results, we have to select the active space wisely, employing some chemical intuition. A summary of the criteria for this selection can be found in [25].

1.3.1 DMRG-CASSCF

As the heart of the CASSCF method is the exponentially scaling FCI method, the limit to the size of the active space, as to the extent of what is still computationally feasible, is roughly CAS(16,16). As many systems require larger active spaces, the reasonable step would be to substitute the FCI method with some approximation. One of the more common choices would be the use of the DMRG method. This leads to the method called DMRG-CASSCF.

1.4 Coupled cluster methods

The Coupled Cluster (CC) method arises from the need to circumvent the problem with truncated-CI methods, where truncating the the FCI expansion at a fixed excitation level leads to a loss of size consistency. This is circumvented by approximating the coefficients of the higher excitations, not included into the expansion, as a combination of products the lower excitation operators.

A convenient notation is using the excitation operator $\hat{T}$, also defined as

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots,$$

where for example the single excitation operator $\hat{T}_1$ in the second quantization form would be

$$\hat{T}_1 = \sum_{a,i} t_i^a \hat{a}_a^\dagger \hat{a}_i,$$

where $a, a^\dagger$ are creation and annihilation operators, and t_i^a are the unknown coefficients of the excitations to be determined, also called the cluster amplitudes. The higher excitation operators are formed in a similar manner.

The ansatz of the wave function is the exponential of the excitation operator applied to the Hartree-Fock solution.

$$|\psi\rangle = \mathrm{e}^{\hat{T}} |\psi_0\rangle$$

This is then substituted into the Schrödinger equation, leading to

$$\hat{H}\mathrm{e}^{\hat{T}}\left|\psi_0\right\rangle = E\mathrm{e}^{\hat{T}}\left|\psi_0\right\rangle, \tag{1.24}$$

or more conveniently

$$\mathrm{e}^{-\hat{T}}\hat{H}\mathrm{e}^{\hat{T}}\left|\psi_0\right\rangle = E\left|\psi_0\right\rangle \tag{1.25}$$

So far, we have not included any approximation, thus, we would not have any computational advantage over FCI. Therefore the next step is to restrict the $\hat{T}$ operator to only some of the excitations, similarly to the truncated CI methods.

To find the solution, we multiply the equation 1.25 from the left by the ground state and different excited wave functions $\langle\psi_0|, \langle\psi_i^a|, \langle\psi_{ij}^{ab}|$ etc., leading to the equations

$$\begin{aligned}
\langle\psi_0|\exp(-\hat{T})\hat{H}\exp\hat{T}\left|\psi_0\right\rangle &= E \\
\langle\psi_i^a|\exp(-\hat{T})\hat{H}\exp\hat{T}\left|\psi_0\right\rangle &= 0 \\
\langle\psi_{ij}^{ab}|\exp(-\hat{T})\hat{H}\exp\hat{T}\left|\psi_0\right\rangle &= 0
\end{aligned}$$

These algebraic equations for the unknown t coefficients and the energy E are nonlinear and usually solved in an iterative procedure.

Most often used are the Coupled Clusters with single and double excitations (CCSD), sometimes with triples included pertubatively (CCSD(T)).

A very detailed description can be found in the book [26].

1.4.1 Tailored coupled clusters

The CC method, as introduced above, is very useful for including dynamical correlation, it usually has much less accuracy in systems with a multireference character. While this could be in principle compensated by increasing the level of excitations, this would cause a dramatical increase of the computational cost.

To overcome this issue, we can use the result of a multireference calculation as a starting point instead. This leads to the class of methods known as Multireference Coupled Clusters (MRCC)[27, 28]. As in this work we focus on the DMRG method, we will consider only the DMRG-Tailored Coupled Cluster Method, accounting for dynamical correlation on top of a DMRG calculation. This is done easily by separating the excitation operator into two parts

$$\hat{T} = \hat{T}_{\mathrm{CAS}} + \hat{T}_{\mathrm{ext}}, \tag{1.26}$$

where the operator $\hat{T}_{\mathrm{CAS}}$ includes all the excitations within the active space, while the $\hat{T}_{\mathrm{ext}}$ groups all the others. The idea is to use the $\hat{T}_{\mathrm{CAS}}$ to account for the static correlation within the active space, while having $\hat{T}_{\mathrm{ext}}$ account for the dynamic correlation, similarly as in the canonical CC method. Both of the operators are truncated to a maximum number of excitations and the resulting cluster operator is used in the exponential ansatz, from where the procedure to find the equations is nearly identical to the one for the canonical CC. The important difference is that for the cluster operator $\hat{T}_{\mathrm{CAS}}$ the amplitudes t_i^a are extracted from the previous multireference calculation. For the DMRG method,

this can be done efficiently, due to the nature of the MPS ansatz and by using the known analytical relations between the cluster amplitudes and the CI coefficients

$$t_i^a = \frac{c_i^a}{c_0}, \tag{1.27a}$$

$$t_{ij}^{ab} = \frac{c_{ij}^{ab}}{c_0} - \frac{c_i^a c_j^b - c_j^a c_i^b}{c_0^2}, \tag{1.27b}$$

for the amplitudes with all indices within the active space orbitals. The $\hat{T}_{\mathrm{CAS}}$ amplitudes are kept constant during the solution of the TCC equations, and only the $\hat{T}_{\mathrm{ext}}$ external amplitudes are optimized. While the method is still formulated using the Hartree-Fock determinant reference, it was shown to be a powerful tool which is able to describe both static and dynamic correlation in a balanced way [29].

1.4.2 LPNO & DLPNO approximations

Although the CC methods are very accurate, for larger systems and/or basis sets one very early gets to their limit due to their prohibitive scaling. To go beyond this limit, the Local Pair Natural Orbitals (LPNO)[30] and Domain based Local Pair Natural Orbital (DLPNO)[31] approximations can be employed, which greatly reduce the computational cost, mainly by exploiting the local nature of electronic correlation by introducing the pair natural orbitals (PNOs)[32]. In the case of the DLPNO approximation, the method even exhibits almost linear scaling with respect to the system size, due to carefully avoiding all worse scaling steps in the procedure.

The procedure proceeds in the following manner:

1. The occupied orbitals are localized.
2. Based on the MP2 energy contributions of pairs of occupied orbitals, some of the pairs are excluded from further TCC treatment and accounted for by only the MP2 method.
3. The pair natural orbitals (PNOs) for each active occupied orbital pair are generated, with carefully prescreening the virtual orbitals to preserve linear scaling.
4. The PNO expansions are then truncated based on their occupation numbers.
5. The resulting PNOs are used in the CCSD or CCSD(T) methods.

This procedure can be easily applied to the TCC methods, as the DMRG-TCC method used in this work [33].

1.5 Adiabatic connection method

While the adiabatic connection formalism was long used within the realm of DFT [34], recently it has been found it can be also applied to a CASSCF wave

function. This led to a novel way to account for the dynamic correlation on top of a multireference calculation [35].

The first step in deriving the adiabatic connection formalism is to introduce a parameter α into the Hamiltonian

$$\hat{H}(\alpha) = \hat{H}_0 + \alpha\hat{H}', \tag{1.28}$$

where $\hat{H}_0$ is the so called group Hamiltonian as defined in [36], in the case of a CASSCF reference wave function having only two terms corresponding to the active and inactive orbitals respectively, and similarly to perturbation theories, $\hat{H}' = \hat{H} - \hat{H}_0$ is the correction. Also, we can formally write the solution

$$\hat{H}(\alpha)\,|\psi(\alpha)\rangle = E(\alpha)\,|\psi(\alpha)\rangle \tag{1.29}$$

This way, formally, we can continuously move from the uncorrelated result ($\alpha = 0$), to the exact solution ($\alpha = 1$).

By using the Hellmann-Feynman theorem and assuming that the one electron density matrix $\gamma(\alpha)$ stays constant with respect to the change of α, an assumption which is valid if the major part of the static correlation is already captured in the CASSCF wave function, the correlation can be then transformed into the form

$$E_{corr}^{\mathrm{AC}} = \int_0^1 W(\alpha)\,\mathrm{d}\alpha, \tag{1.30}$$

with the integrand $W(\alpha)$ expressible using the Transition Density Matrices (TDMs) and the natural occupancies of the active space orbitals. To obtain the value of the correlation energy, we employ the Extended Random Phase Approximation (ERPA) for the TDMs and numerically evaluate the integral in equation 1.30.

1.5.1 AC0

The AC method can be further simplified by expanding the integrand in the AC energy expression 1.30 into a Taylor series at $\alpha = 0$ and keeping only the linear term, as the constant term vanishes for a CASSCF reference. Introducing the notation

$$W^{(1)} = \left.\frac{\partial W(\alpha)}{\partial\alpha}\right|_{\alpha=0},$$

and evaluating the linearized integral leads us to the energy expression

$$E_{corr}^{\mathrm{AC0}} = \frac{W^{(1)}}{2} \tag{1.31}$$

The $W^{(1)}$ term has an analytic form and is expressed, after applying the ERPA, from the 1 and 2-particle reduced density matrices (RDMs)[36]. This surpasses the need for the expensive integral evaluation, significantly reducing the computational cost.

Compared with NEVPT2, one of the most used methods for multireference dynamic correlation treatment, it achieves a similar or greater accuracy with a greatly reduced computational cost, especially for larger active spaces [37]. This is due the AC only using up to the 2-electron RDMs, in contrast with NEVPT2 requiring also the 3 and 4- electron RDMs, which leads to a prohibitive scaling with the active space size of the NEVPT2 method.

2. Studied systems

2.1 Porphyrin model

The first studied system was a model of the porphyrin molecule, which was taken from the work of Li Manni et al.[9]. This model uses the geometry of the Fe(II)-porphyrin, simplified by removing the 2-carbons most distant from the metal center on each of the 4 pyrrole rings and replacing them by hydrogen atoms. This not only reduces the number of basis set functions, but more importantly, reduces the needed active space for multireference treatment by eliminating 8 of the porphyrin π-orbitals. The geometry is shown in the Figure 2.1.

The primary motivation for the study of this system was a discrepancy between the experimental predictions and quantum chemistry calculations for the spin of the ground state. The Mössbauer spectroscopy experiments [38], which were an inspiration for these calculations, show that a similar molecule, Fe(II)-phtalocyanine (figure 2.2), exhibits a triplet ground state in crystalline form or in a polar solvent, while in a non-polar solvent, which most closely resembles the vacuum conditions of quantum chemistry, it adopts a quintet ground state. This is accompanied by a significant change in geometry, with the most notable difference being the Fe-N distance, which grows from 1.926 Å in the crystalline state to > 2 Å while isolated. Unfortunately, as the nuclear positions can be successfully measured only in the crystalline state by X-ray diffraction, we still have to rely on quantum chemistry to find the geometry of the free molecule. Similarly, for the Fe(II)-tetraphenylporphyrin (Fe-TPP) molecule (figure 2.2), current experimental data [39] suggest a triplet ground state. But as the experiments again have been performed in the crystalline form or polar solvents, this could cause artificial stabilization of the triplet state, similarly to the Fe(II)-phtalocyanine molecule. The effect of the environment would explain the discrepancy between the experiment and computational results, which (mostly) predict the quintet ground state.

Another motivation was to compare the performance of the recently developed tailored coupled cluster methods [29, 33] with other methods applied on the porphyrin system.

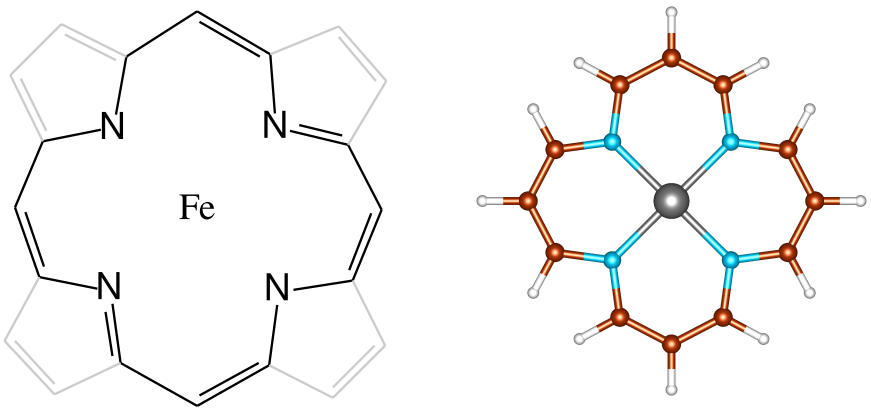

Figure 2.1: The geometry of the porphyrin model.

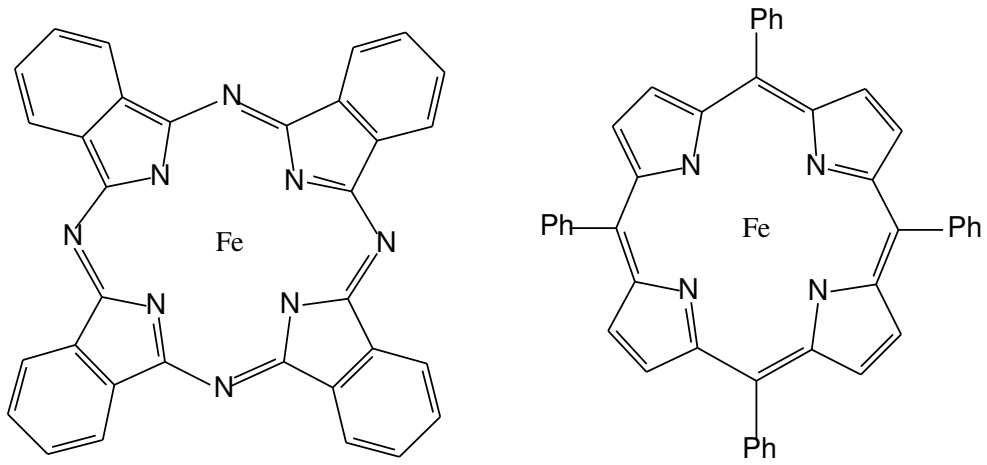

Figure 2.2: The Fe-phtalocyanin (left) and Fe-TPP (right) molecules.

The previously mentioned experimental results lead to the assumption that geometrical changes are important for the spin chemistry of porphyrin systems and using a single geometry for different spin states would introduce a bias into the results. It also indicates we have to study the adiabatic spin gap if we want to correctly describe the behavior of these systems.

In order to do this, we have used three different, but very similar geometries, their most significant difference being the Fe-N distance. One of them, for reference, was taken from the mentioned work of Li Manni et al.[9] which they obtained by simplifying the geometry of the full porphyrin from the study of Pierloot [40], while the other two were fully optimized at the B97-D3/def2-TZVPP level, for the triplet and quintet spin states. The Fe-N distances were 1.989Å for the Li Manni reference geometry, 2.048Å for the triplet optimized geometry and 2.180Å for the quintet optimized one.

For each of these three different geometries, we have performed two separate (DMRG-)CASSCF calculations for the triplet and quintet states to account for the static correlation, which we then used for subsequent TCC calculations to include the dynamic correlation.

We have used three different active space sizes.

- CAS(8,12) containing the 3d orbitals, the double shell 4d orbitals, the Fe σ-bonding orbital and the Fe 4s orbital.
- CAS(12,16), which we got by adding two occupied and two unoccupied π-orbitals of the porphyrin to the previous active space. While they were chosen by their single orbital entropies, they correspond precisely to the orbitals of the Gouterman model [41].
- The full CAS(32,34), containing the whole π-system of the porphyrin, the Fe 3d, 4s, 4p & 4d orbitals and the σ-bonding orbitals between the iron and porphyrin.

The first two active spaces were small enough to be done using conventional CASSCF, for the third one we have used DMRG-CASSCF. We have used a fixed bond dimension of 1024 for the DMRG calculations.

Using the CASSCF orbitals, we performed a single point DMRG calculation using the DBSS approach with a truncation error of 10^{-6} to get the amplitudes for subsequent TCCSD and TCCSD(T) calculations.

We used two basis sets, def2-SVP and def2-TZVP [42]. The smaller one enabled us to use the canonical TCC method, while the larger one could be done only under the DLPNO [31] approximation. In the smaller basis set, we were able to compare the canonical and DLPNO-TCC assessing the accuracy of the the DLPNO approximation.

The geometry optimizations were done using the Gaussian [43] program, while the CASSCF and DMRG-CASSCF were done using the ORCA [44] program package, interfaced with the Budapest QC-DMRG program [45].

The DMRG calculations for the TCC amplitudes were done using the new MOLMPS package [46], while the TCC calculations themselves were done again using the ORCA package.

Some of the results also motivated us to do a calculation on the same model of the porphine molecule. The geometry was obtained by removing the central iron from the geometry of Li-Manni, adding two hydrogen atoms, and optimizing the positions of the two hydrogens, while fixing the rest of the molecule. The geometry optimalization was again carried out at the B97-D3/def2-TZVPP level. We have then continued with a CASSCF calculation of the full π-system of the molecule in the def2-TZVP basis set. This resulted in a CAS(18,16).

2.2 Iron porphyrin carbene system

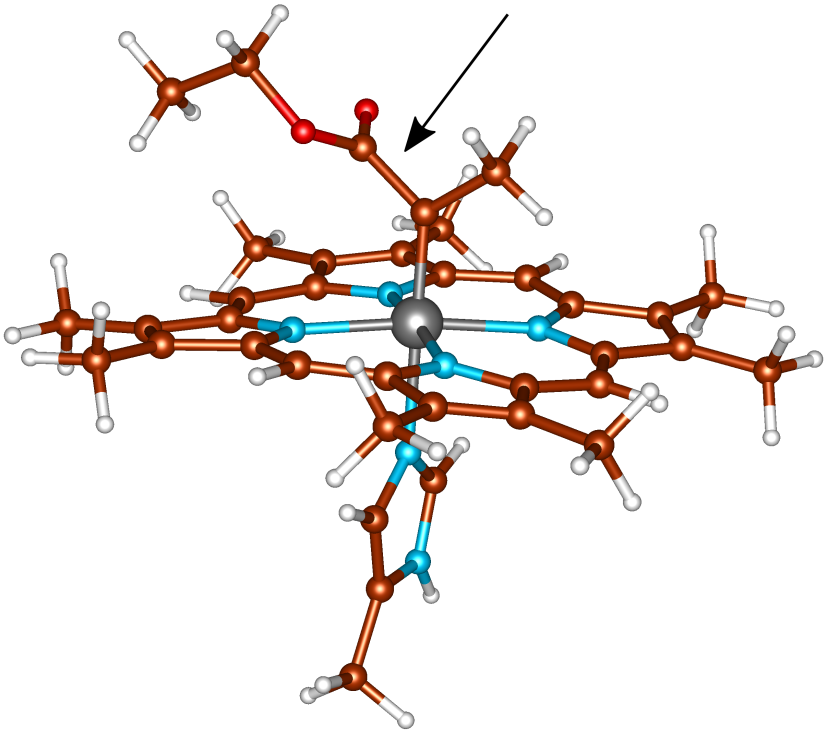

Figure 2.3: One of the geometries used in our calculations. The bond, around which the torsion was studied, is indicated by the arrow.

To use the obtained accurate results on a larger system with potential applications, we studied a Fe(II) porphyrin molecule bound axially to a carbene ligand. These porphyrin ligand systems, known in literature as iron porphyrin carbenes (IPCs), are interesting, as they are an important intermediale in reactions involving a transfer of the carbene group, using both some wild-type and engineered enzymes [47]. While these reactions are not found in nature, some of the wild enzymes of the cytochrome group proved to be useful for these reactions, followed by engineering artificial enzymes better suited to meet specific needs. They can be used for many reactions, including some which insert various heteroatoms like boron, nitrogen, silicon or sulfur. Also, given by the chirality of the catalyst, the reactions can be made stereoselective.

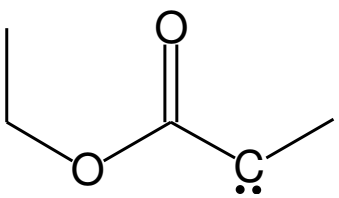

Figure 2.4: The chemical structure of the carbene resulting from the decomposition of ethyl 2-diazopropanoate, which was used as the axial ligand in our calculations.

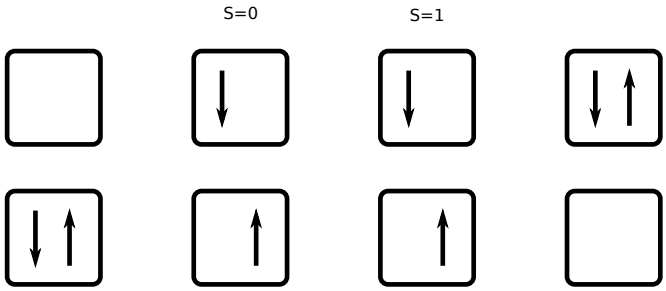

Figure 2.5: The full model space CAS(2,2) we were trying to find, in order to use the ful model space approach. Note that the middle two determinants mean a spin adapted linear combination instead.

For these calculations we used the geometry of the full Fe-porphyrin, with added methyl groups in the β positions supplementing all of the sidechains of the original heme molecule. Axially bound to the central iron were an imidazole molecule, simulating the histidine of the original protein binding the porphyrin in the active site, and the carbene resulting from the decomposition of ethyl 2-diazopropanoate, representing one of the possible substrates for the enzyme. The geometry was optimized using the B3LYP/6-311G(d) level. One of the geometries used, in the middle of the torsion angle, is shown in the Figure 2.3. The chemical structure of the ligand is represented in figure 2.4.

Introducing the two axially bound ligands changes the occupation of the d-orbitals of the central iron, causing the lowest spin states be singlets and triplets. We decided to include the two lowest singlet and triplet states into the calculations. We were again interested in the relative energies of these states and also the change caused by rotating along the torsion angle of one of the bonds on the carbene ligand.

As for this system we were interested also in excited states, the basic TCC treatment of the dynamic correlation was not possible, as it allows only ground state calculations for each of the spin states. Thus, we attempted to proceed using a Hilbert space multireference generalization of TCC using the full model space approach. For this we had to select 3 excited states, which together with the ground state would comprise the active space CAS(2,2). This active space is schematically shown in figure 2.5. The lowest excited states are in the attachment A.3.

Unfortunately, the electronic structure of the system was found to be very complicated. This led to the doubly excited state not being among the lowest lying electronic states accessible by the DMRG method. While this could be circumvented by several approaches within the coupled cluster methodology, they haven't been implemented yet in our group. Therefore we decided to test an alternative by using the AC0 method which is not limited by this issue. To control the accuracy of the AC0 method we performed also NEVPT2 calculations, which we could directly compare.

The use of AC0 also had another motivation. Based on previous studies, the method has an accuracy similar to NEVPT2 [36]. But, as was mentioned earlier, NEVPT2 uses the 3 and 4 electron reduced density matrices (RDMs), while AC0 uses only 2 electron RDMs, which means AC0 scales much better with respect

to increasing the active space size. This is crucial when transitioning to DMRG-CASSCF, that is intended for use with large active spaces, beyond the limits of conventional CASSCF. Thus, we hoped to compare the AC0 and NEVPT2 energies to see whether the AC0 gives reasonable results and is computationally feasible for larger systems.

Due to the molecule being much larger than the previous model, we have not used DMRG-CASSCF with the full active space, but instead chose a suitable smaller active space, accessible by conventional CASSCF, consisting of the 3d orbitals, the σ-bonding orbital to the ligand, an empty p-orbital on the carbene atom and 5 π-orbitals on the porphyrin, leading to a CAS(14,12).

As the number of basis set functions we were able to use for the AC0 and NEVPT2 calculations was limited, we developed a hybrid basis set, which was chosen to match the results at the CASSCF level in a larger (cc-PVQZ) basis set. The comparison between the basis sets is in the attachement A.4.

As a result we had two basis sets. A smaller one consisting of 516 functions, and a larger one, with 864 functions. The smaller basis set consisted of the 6-31G basis set, with 6-31G(d) functions on the 4 sigma bonded nitrogen atoms in the porphyrin ring and 6-311G(d) functions on both the central iron and the axially bound carbene carbon, while the larger basis was the 6-31G(d) basis set improved by adding the 6-311G(d) functions on the 4 nitrogens, the central iron and the whole carbene ligand.

3. Results

3.1 Porphyrin model

Throughout this work, we have used this convention for the triplet-quintet spin gap

$$\Delta E = E_T - E_Q. \tag{3.1}$$

This convention ensures that the energy value for the spin gap is positive for a quintet ground state and negative for a triplet ground state.

The natural orbitals from the DMRG-CASSCF calculations with the largest CAS(32,34) active space are shown in Figures 3.2, 3.3 for the triplet and quintet states respectively. The DLPNO approximation performs well for this system, as can be seen from Figure 3.1. The error of the spin gap energies compared to the canonical approach is less than 0.5 kcal/mol, except for one of the geometries using the smallest active space, where it was equal to 0.6 kcal/mol. This is less than the expected errors caused by the size of the used basis sets. It is reasonable to expect the calculations in the def2-TZVP basis set will be affected by errors of similar (marginal) size.

The calculations on the reference geometry yielded similar results to those in the original work [9], which used stochastic-CASSCF for the largest active space calculations, as can be seen from Figure 3.1. There was a difference between the used basis sets, which despite being of approximately the same size, still could have caused a non-negligible error. Another difference is that we haven't included the DKH relativistic corrections. These inconsistencies not only led to slightly different sizes of the quintet triplet spin gap, but also led to the lowest triplet state being ${}^3A_{2g}$ with the occupation on the iron being $(d_{x^2-y^2})^2(d_{z^2})^2(d_{xz})^1(d_{yz})^1(d_{xy})^0$ unlike the ${}^3A_{2g}$ state in ref. [9]. On the other hand, the two states in question are nearly degenerate with the difference being < 1 kcal/mol at the CASSCF level, both in the mentioned work and our calculations. The separation of energies for the smallest two active spaces in the reference, not present in our calculations, is most probably caused by a difference in the choice of the medium sized active space with the reference including the 3s instead of the 4s orbital, obtaining a CAS(14,16) instead.

The resulting vertical spin gaps for the calculations on the reference geometry in the TZVP basis set are presented in the Figure 3.1. There we see a huge difference at the CASSCF level between the first two used active spaces which have a similar value for the spin gap, and the largest one where the spin ordering reverses.

Table 3.1: The error of the DLPNO approximation compared to the canonical TCCSD and TCCSD(T) in the def2-SVP basis set. The energies are in kcal/mol

	CAS(8,12)		CAS(12,16)		CAS(32,34)	
	TCCSD	TCCSD(T)	TCCSD	TCCSD(T)	TCCSD	TCCSD(T)
Reference geometry	0.59	0.65	0.23	0.27	-0.09	-0.06
Triplet geometry	0.05	0.08	-0.05	-0.04	-0.20	-0.25
Quintet geometry	-0.07	-0.04	-0.06	-0.06	-0.03	-0.06

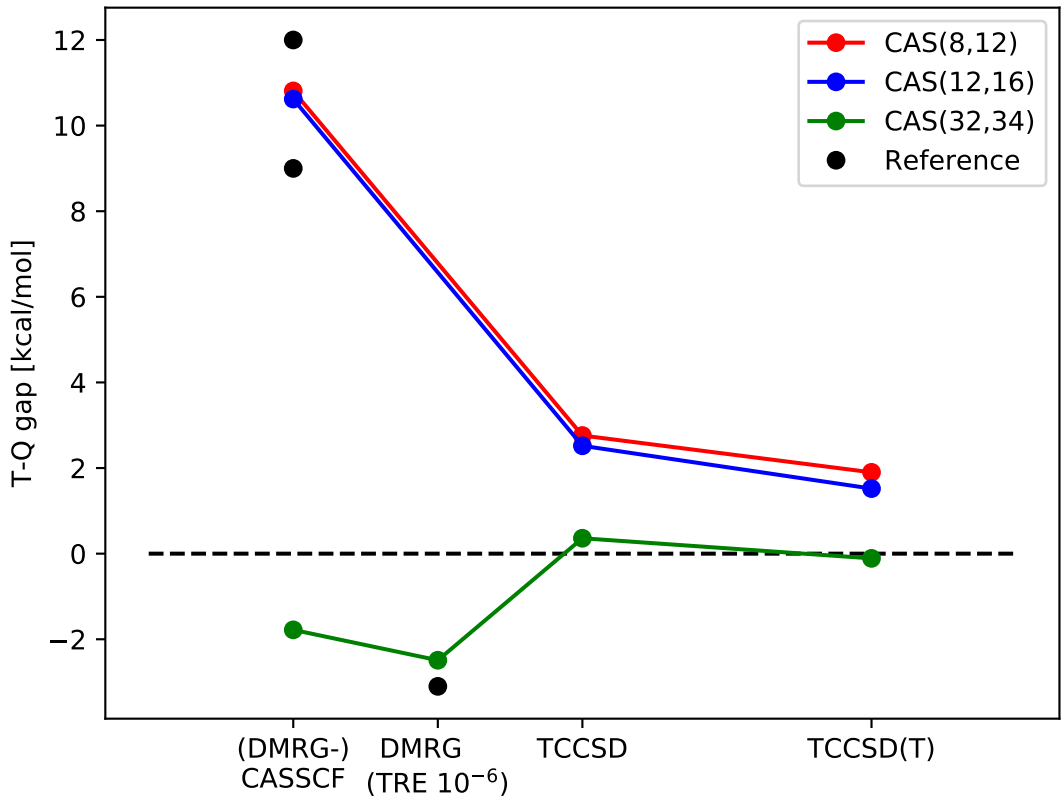

Figure 3.1: The vertical quintet-triplet gap on the reference geometry in the def2-TZVP basis set. A positive spin gap corresponds to a quintet ground state. The reference data is taken from [9]

On the other hand, including the dynamical correlation reduces the difference, as it makes up for the missing π-system correlation, although inadequately, thus still introducing a significant error of ~ 2 kcal/mol.

So, while the dynamic correlation for the smaller active spaces stabilizes the triplet state, this is mostly due to compensating for the missing static correlation in the porphyrin π-system, and the dynamic correlation in fact stabilizes the quintet state, as can be seen from the CAS(32,34) calculations.

Surprisingly, the inclusion of the 4 π-orbitals had hardly any effect on the spin gap, which is unexpected, as those orbitals are the orbitals in the Gouterman model, which explains the visible spectra of porphyrin molecules by the means of excitations between these orbitals. On the other hand, while at the CASSCF level the difference was insignificant for all three geometries, for some of the geometries there is a slightly more significant effect at the level of dynamical correlation treatment.

In Figure 3.4 we see that the value of the spin gap is strongly dependent on the used geometry, where increasing the Fe-N distance enlarges the gap. While the dependence seems to be roughly linear, this is probably just due to the small range of the studied Fe-N distances. The stabilization of the quintet state by the larger Fe-N separation is predictable, as asymptotically the Fe^{2+} cation is in the quintet state, while the free porphine is a singlet. This, given the standard rules for the combination of quantized angular momentum, leads to a quintet ground

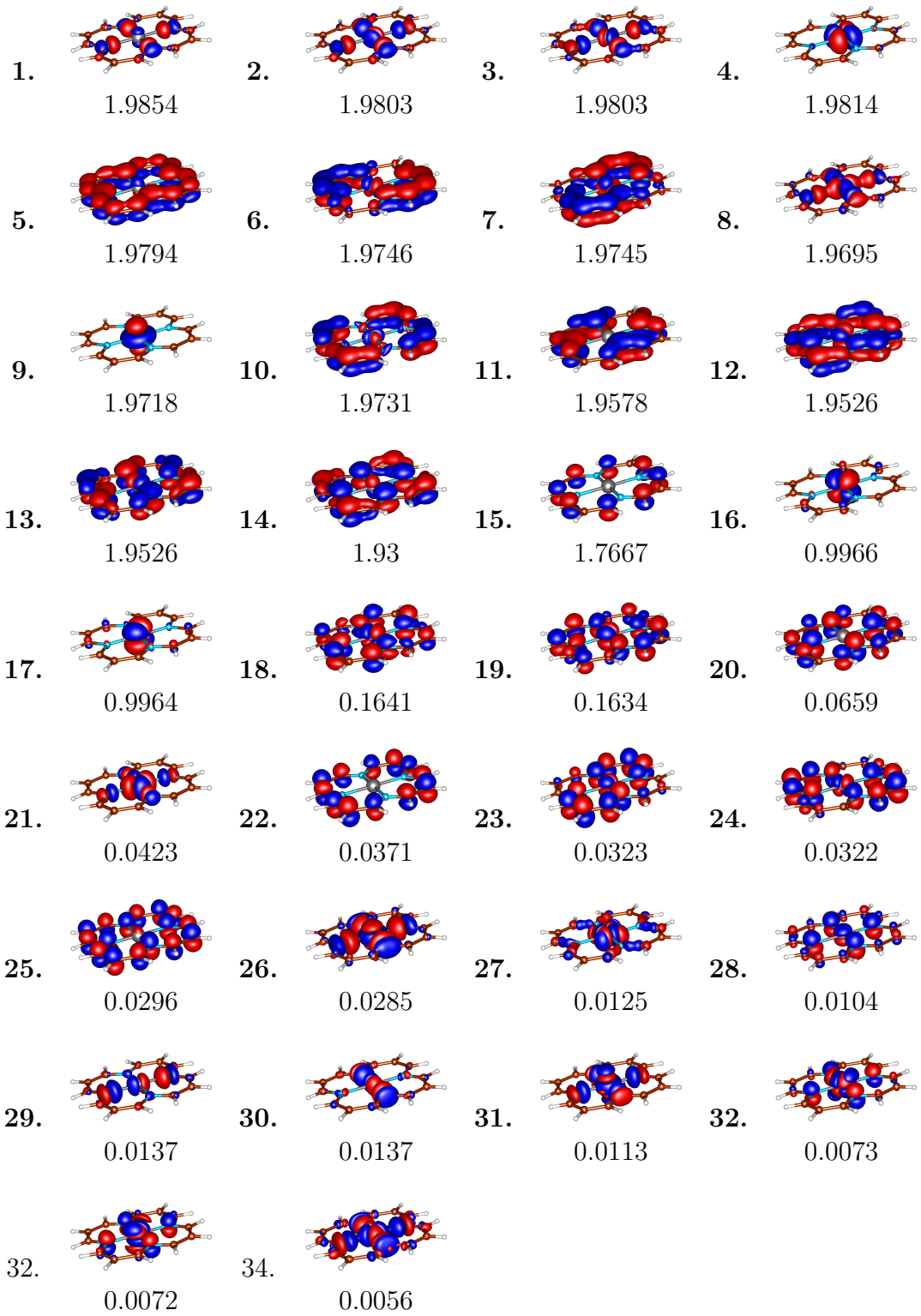

Figure 3.2: The DMRG-CASSCF natural orbitals from the triplet calculations in the TZVP basis set. The numbers under the images represent the natural occupation numbers.

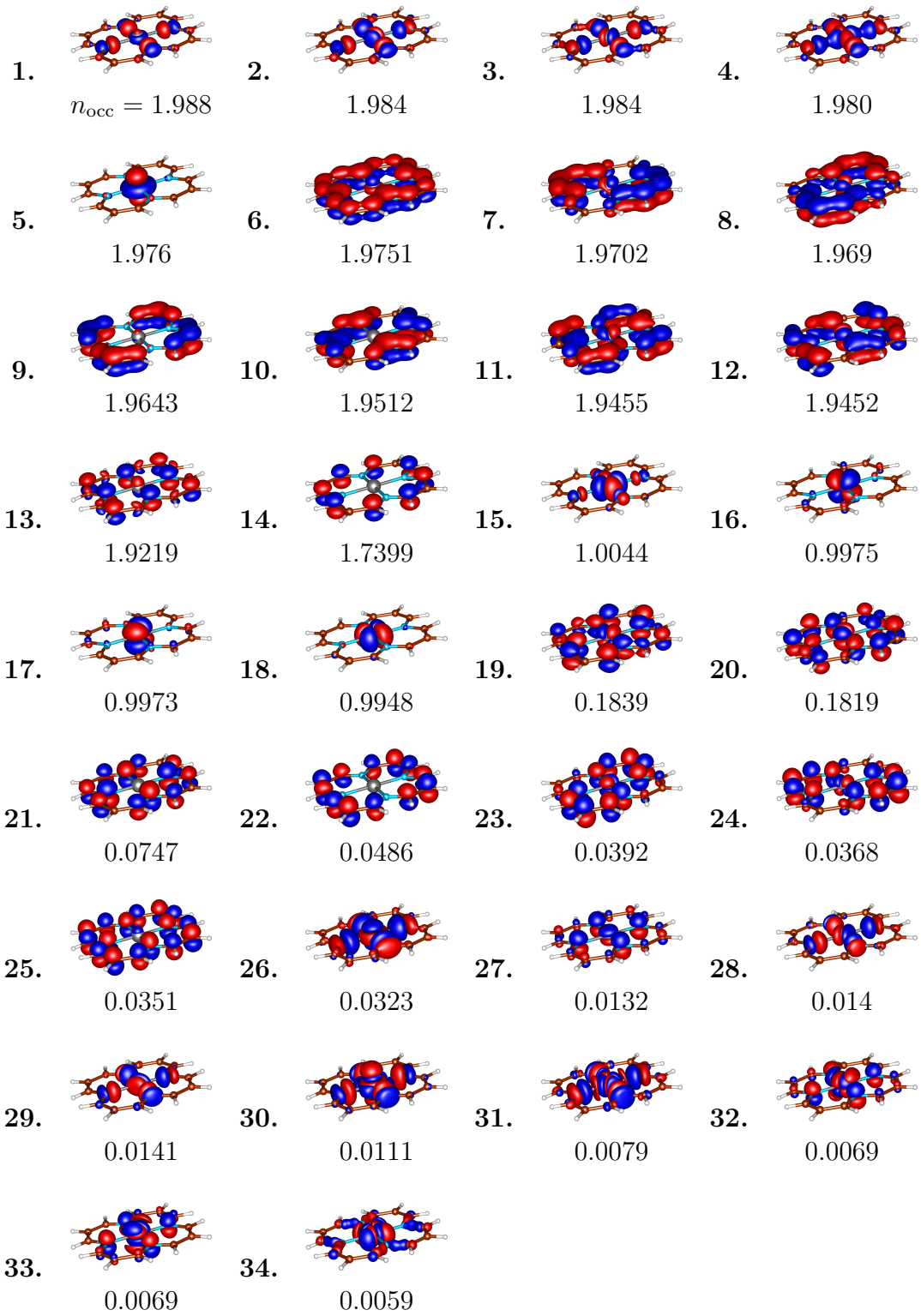

Figure 3.3: The DMRG-CASSCF natural orbitals from the quintet calculations in the TZVP basis set. The numbers under the images represent the natural occupation numbers

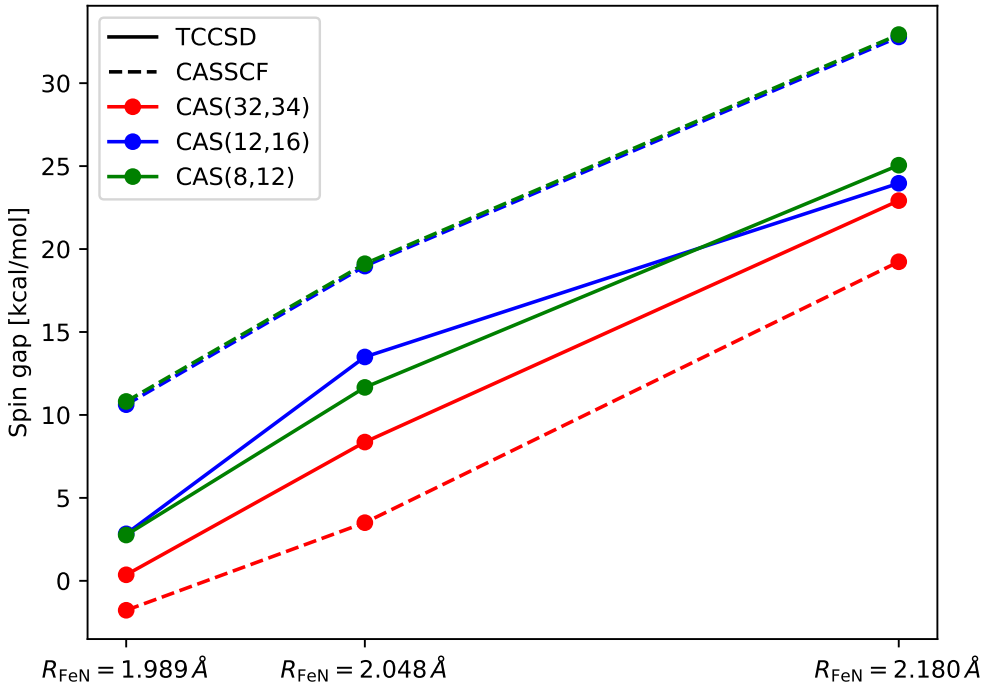

Figure 3.4: The effect of geometry on the value of the vertical triplet-quintet spin gap, at the CASSCF and DLPNO-TCCSD levels of theory in the TZVP basis set.

state in the limit of no iron-porphyrin interaction. This means that the triplet state must be stabilized by the iron-porphyrin interaction.

However, the full porphyrin molecule has much more rigidity compared to our model. This is due to the porphyrin having the full pyrrole rings, causing sterical strain that does not allow as prominent geometrical changes compared to our model. This is also seen in comparing the Fe-N distances of our model and the optimal geometries for the full porphyrin, which are shown in Table 3.2.

Table 3.2: The Fe-N distances for the geometries of the porphyrin molecule and model, optimized at the B97-D3/def2-TZVPP level for both the triplet and quintet spin states, except for the reference one, which was taken from Ref. [9].

		Triplet	Quintet
Model	- reference	1.989	
	- optimized	2.048	2.180
Fe(II)-Porphyrin		1.997	2.064

By comparing the Fe-N distances of the geometries of the model to those of the full porphyrin we see, that due to increased flexibility, the Fe-N distance of the triplet state of the model approximately corresponds to the quintet state of the full porphyrin. This means that the calculations on the triplet optimized

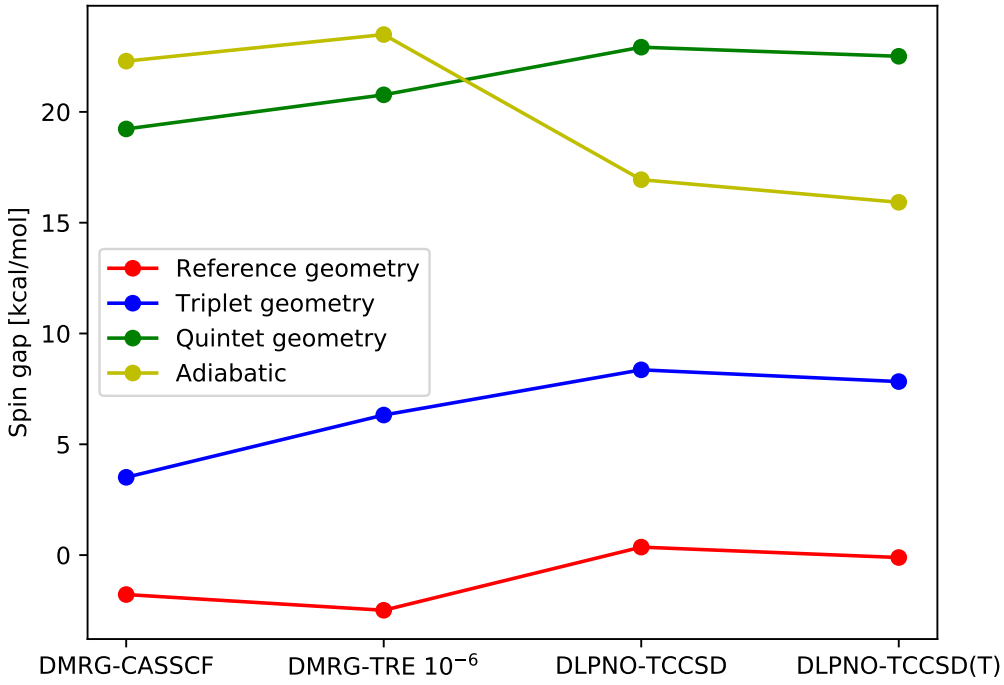

Figure 3.5: The triplet-quintet vertical and adiabatic spin gap at the CAS(32,34) level. The adiabatic gap is the difference between the triplet state of the triplet geometry and the quintet state of the quintet geometry

geometry should reflect the vertical excitation energies of the quintet state of the full Fe-porphyrin much better than the triplet state.

The resulting adiabatic triplet-quintet spin gap, together with the vertical ones, all of them calculated using the CAS(32,34) in the TZVP basis set, are shown in Figure 3.5. We see that considering the adiabatic spin gap greatly stabilizes the quintet. However, the missing energy of the sterical strain causes the adiabatic spin gap to overshoot compared to the full porphyrin, thus undermining further discussions of the adiabatic spin gap of the Fe(II)-porphyrin, based on this model.

To gain further insights, we performed calculations on a model of the porphine molecule without the central iron. The resulting singlet-triplet gap of the porphine model was 8.22 kcal/mol. Comparing this with the experimental value of 36.44 kcal/mol for the porphyrin molecule taken from [48] and the computational value of 37.36 kcal/mol [49], we see that the introduced geometrical approximations, which are quite drastic, cause a large change in the value of the spin gap of the porphine model, which naturally will propagate itself into the error of the spin gap of the porphyrin model with iron. However, from the aforementioned discussions of the asymptotic behavior, this would mean the triplet could be favored by the introduced geometrical approximation, which would mean further stabilization of the quintet state in the full porphyrin.

While our model is too crude to provide any quantitative results about the full porphyrin molecule, the state of the art DMRG-CASSCF calculations on the

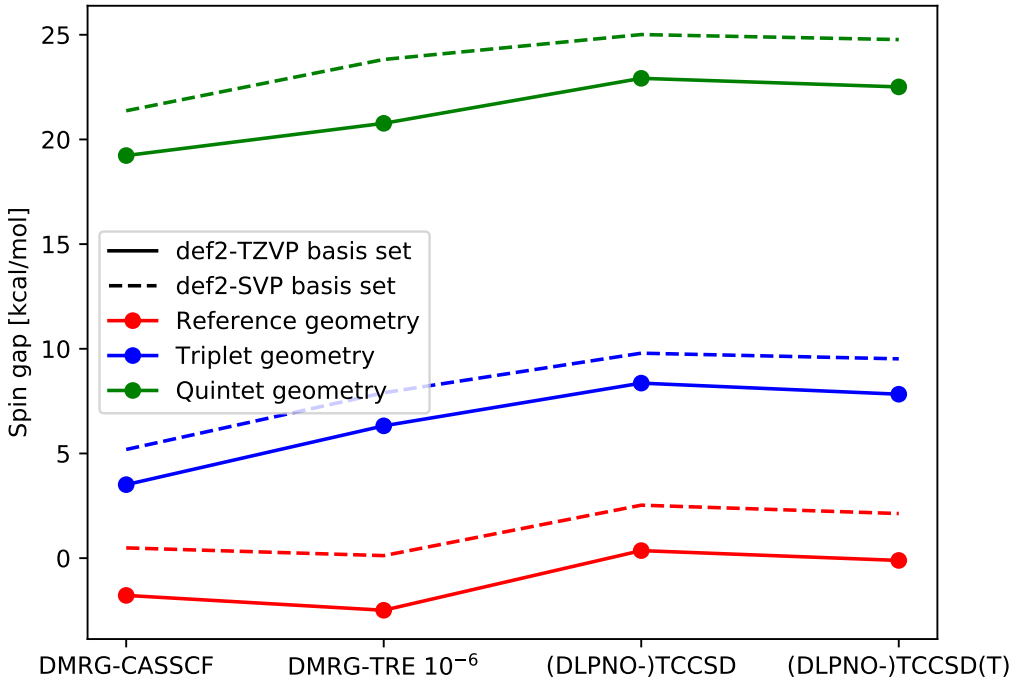

Figure 3.6: The results for the CAS(32,34) calculations for all three geometries, using the two different basis sets. In the def2-SVP basis set we show the result of the canonical TCC approach2, while in the def2-TZVP basis set, we display the DLPNO aproximated TCC results.

model with the geometry with an Fe-N distance similar to that of the triplet state for the full Fe(II)-porphyrin molecule place the triplet as the ground state with a small margin of 2.5 kcal/mol and including the dynamical correlation via the TCCSD(T) method causes the two states to become near degenerate, for a geometry close to that of the quintet state of the full porphyrin we have the quintet ground state well separated by 7.8 kcal/mol at the TCCSD(T) level. Thus, by considering the geometrical changes associated with different spin states, we are still able to predict with enough confidence that the free Fe(II)-porphyrin molecule has a quintet ground state.

This is consistent with the experiments on the Fe(II)-phtalocyanine molecule [38], which undergoes significant geometrical relaxation between the triplet and quintet states. Also, those experimental results indicate that the crystalline state or the environment of a polar solvent stabilizes the triplet spin state of unligated Fe(II)phtalocyanine, which in a non-polar environment has a quintet state. Similarly, for the Fe(II) tetraphenylporphyrine molecule experiments predict a triplet ground state in the crystalline state, which is consistent with the findings of us and [9]. However, due to geometrical relaxation, which may not be allowed in the highly constrained crystalline state, we can still predict the quintet state for the isolated porphyrin molecule.

From the graphs of 1-orbital entropies 3.7 we see that including both four and five π-orbitals into the calculations with the smaller active spaces is justified,

as the 5 orbitals having the highest single orbital entropy are π-orbitals on the porphyrin ring, which are also well separated in that manner from the rest of the active space orbitals, while the first 4 π-orbitals constitute the well-established Gouterman model.

The more interesting thing about the graphs is that the entropy of the d-orbitals on iron is significantly lower than most of the π-orbitals of the porphyrin ring, which shows that in fact, there is significant static correlation within the porphyrin π-orbital space. This can be also seen from the circular mutual information plots 3.8, where we can also see that the total entropy stems mostly from the entanglement of the π-orbitals of the porphyrin and not from the porphyrin-iron interaction. However, although there is no significant mixing on between the π-space and the natural orbitals of a purely d-character on the iron atom, due to proper spatial symmetry, there is d_{xz} and d_{yz} iron orbitals actively mix themselves into some of the π-orbitals. The single orbital entropy values are also consistent with the fact, that the most strongly non-integer natural occupation numbers correspond to some of the π-orbitals on the porphyrin ring, as can be seen in Figures 3.2 and 3.3.

Comparing the single site entropy plots 3.7 and the mutual information graphs 3.8 for the two spin states, we see that the entanglement is slightly higher for the triplet state. This is consistent with the triplet state being stabilized by the iron porphyrin interaction. The higher correlation of the triplet state is consistent with the enlarging the basis set stabilizing the triplet, as the increased flexibility is needed to more adequately describe the correlation.

The static correlation on the porphyrin molecule can best be seen from the previously mentioned calculations, where we performed CASSCF on a similar model of the porphine molecule, by seeing how much removing the iron atom reduces the multireference character of the system. The results indicate that the multireference character is more or less unchanged, shown both by the coefficient of the highest contributing determinant being 0.57 for the singlet ground state of our porphine model, which is actually less then for any of the calculations of the original porphyrin model with the iron included. Also the natural occupation numbers were close to those of the corresponding π-orbitals in the CAS(32,34) calculation.

This implies that single reference treatment is not appropriate if we need accurate results, and porphyrin systems generally call for a multireference treatment, as the π-orbitals mix with some of the d-orbitals because of proper spatial symmetry, as can be seen from the natural orbitals of the CAS(32, 34) calculation.

Also, from both the results on the value of the spin gap of the model shown in Figure 3.1, and the calculations on the porphine molecule, we can see that the full π-electron system of the porphyrin has to be included into the active space, as dynamic correlation treatment cannot fully account for the incomplete active space in terms of correlation energy.

Unfortunately, this means that the necessary active spaces are large, well out of the reach of conventional CASSCF. Therefore, at the moment we are working on an efficient implementation of DMRG-CASSCF, which would allow us to move to larger active spaces, for example the CAS(40,42) of the full Fe(II)-porphyrin.

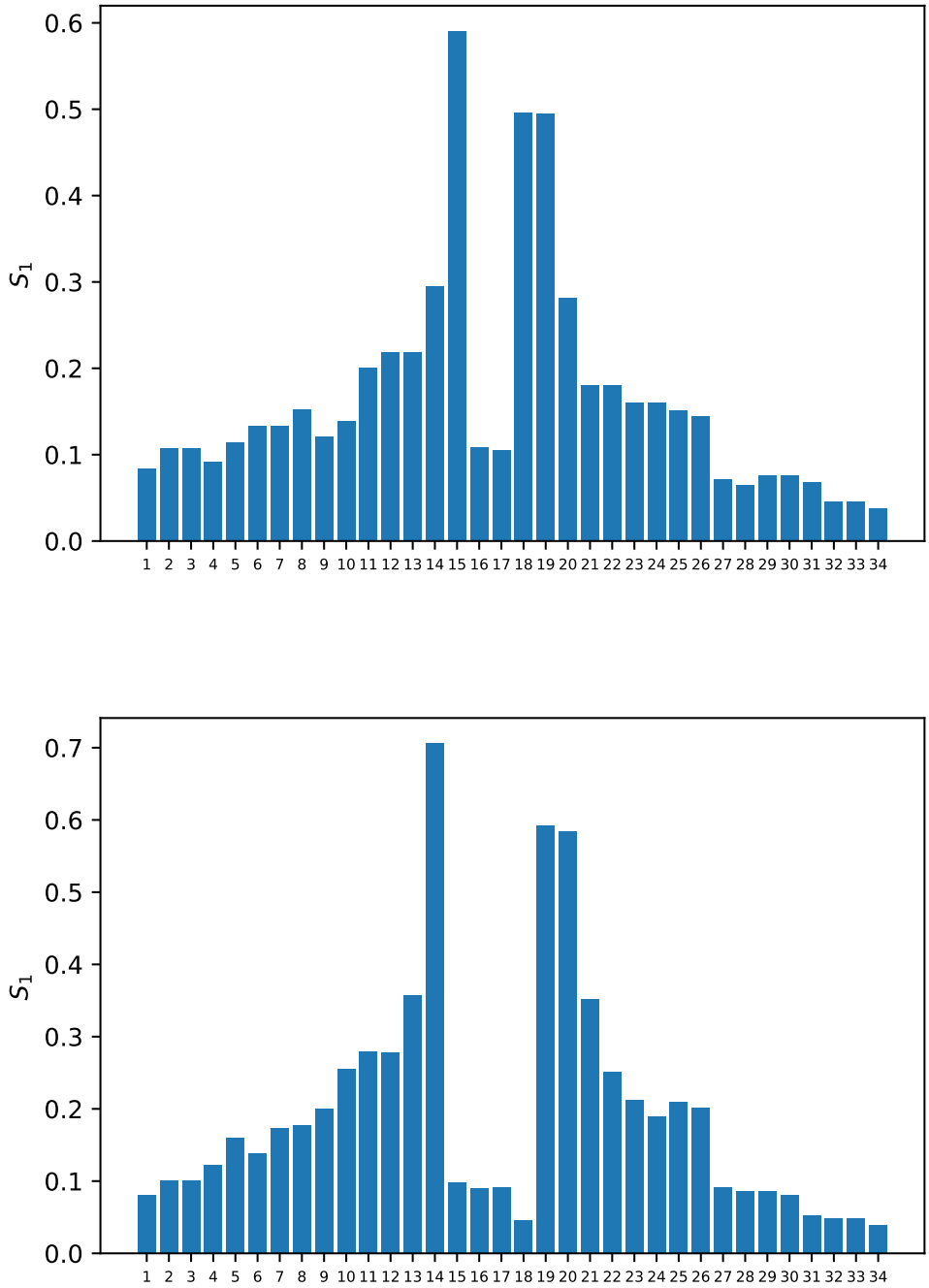

Figure 3.7: Graph of the single orbital entropies for the triplet (top) and quintet (bottom) states on the reference geometry. The orbital labeling is the same as the one of the natural orbitals in the figures 3.2 and 3.3.

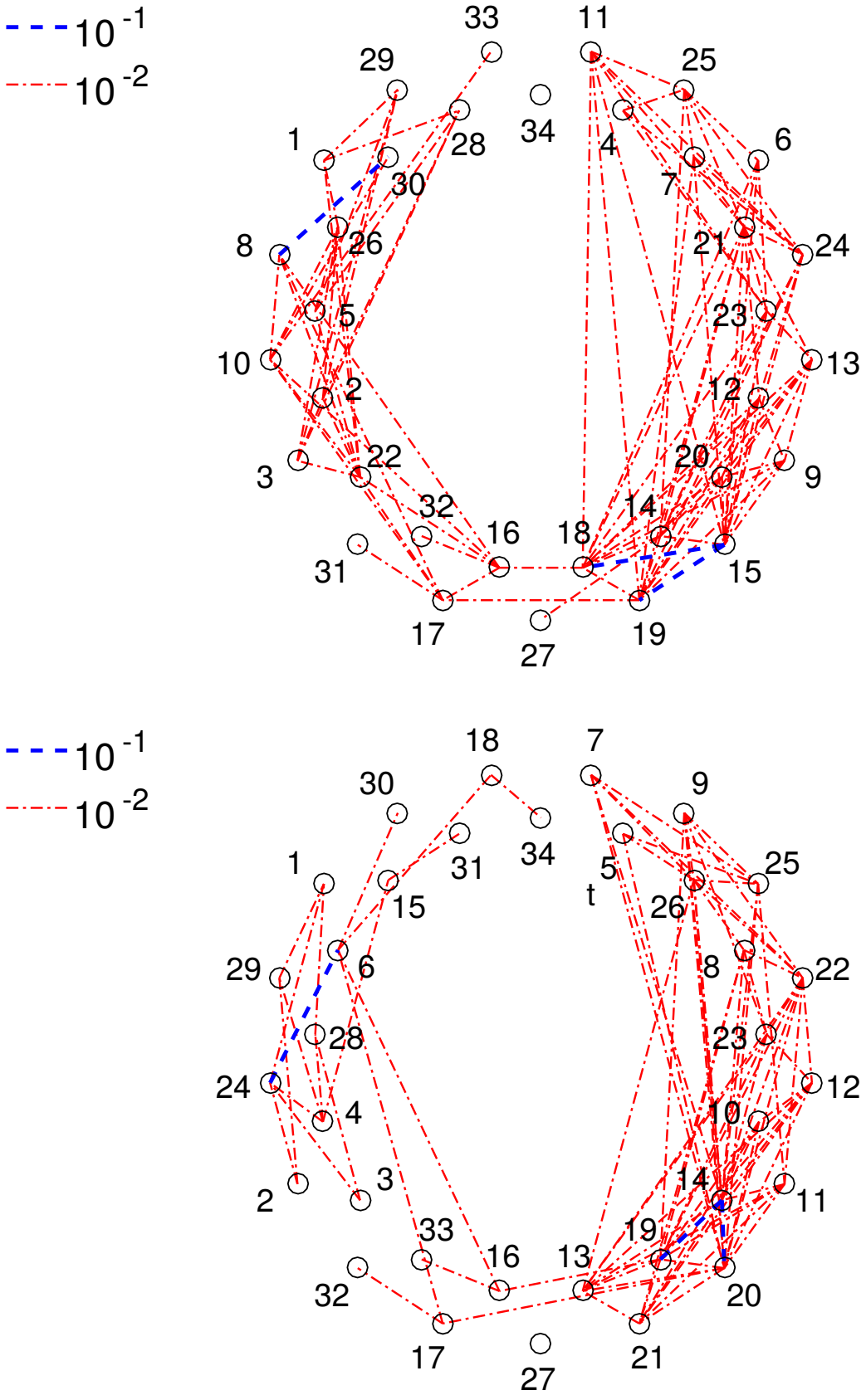

Figure 3.8: The graphs of mutual information for the triplet (top) and the quintet (bottom) states. The orbital labeling is the same as in the figures 3.2, 3.3.

3.2 Iron carbene porphyrin system

The natural orbitals from the SA-CASSCF calculations on the geometry in the middle of the torsion angle curve are shown in the Figure 3.10. The natural occupation numbers are not included, as they have little physical significance for SA calculations, given that they are averaged over the 4 states.

We have tried both the state-averaged CASSCF (SA-CASSCF) and the state specific approach. Unfortunately, for the state specific approach, we weren't able to achieve convergence for some of the states (namely the S_0 and T_1 states), a typical problem associated with state specific calculations. Therefore we only proceeded using the SA-CASSCF approach for the 4 states in the subsequent dynamical correlation treatment calculations.

In the graph 3.9 we have the CASSCF and AC0 results for the relative energies of the 4 spin states along the torsion angle. As we can see, the CASSCF and AC0 do not correspond well. Considering on the results with the CAS(12,16) of the porphyrin model, it is probably mostly due to the AC0 compensating for insufficient treatment of the static correlation on the porphyrin ring, by an incomplete active space. An interesting effect is, that at the CASSCF level, the ground state is near degenerate between the singlet and triplet states, with their relative order switching if the basis set size is changed, but including the dynamic correlation (and also compensating for the missing static one) strongly stabilizes

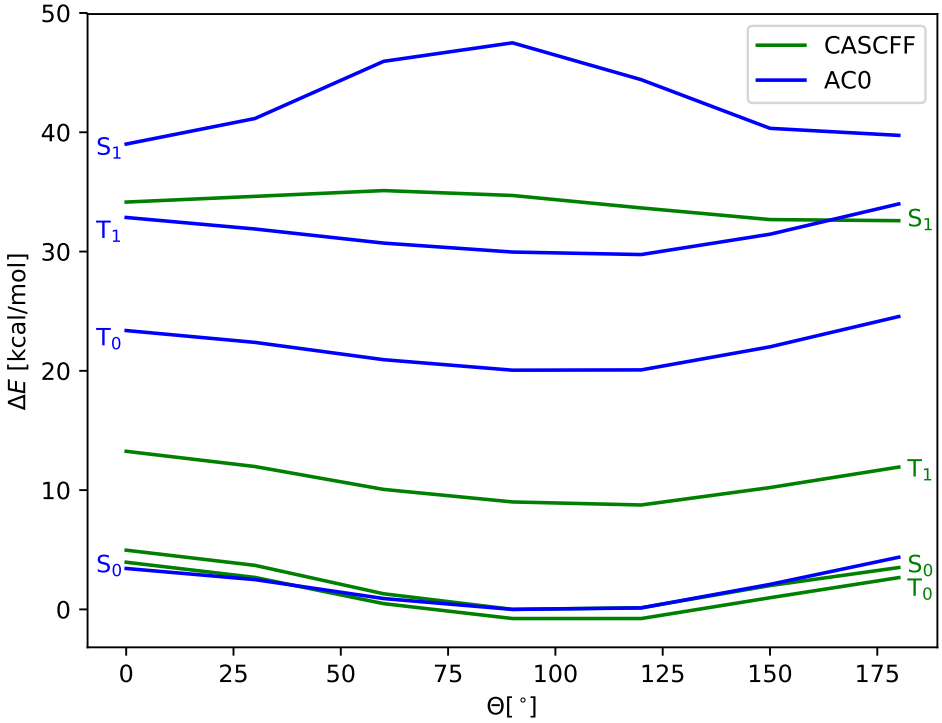

Figure 3.9: The relative energies of the four states along the torsion angle, at the CASSCF and AC0 levels. The energies for all states were shifted so that the energy of lowest singlet state at the middle of the curve is equal to zero. Θ represents the torsion angle.

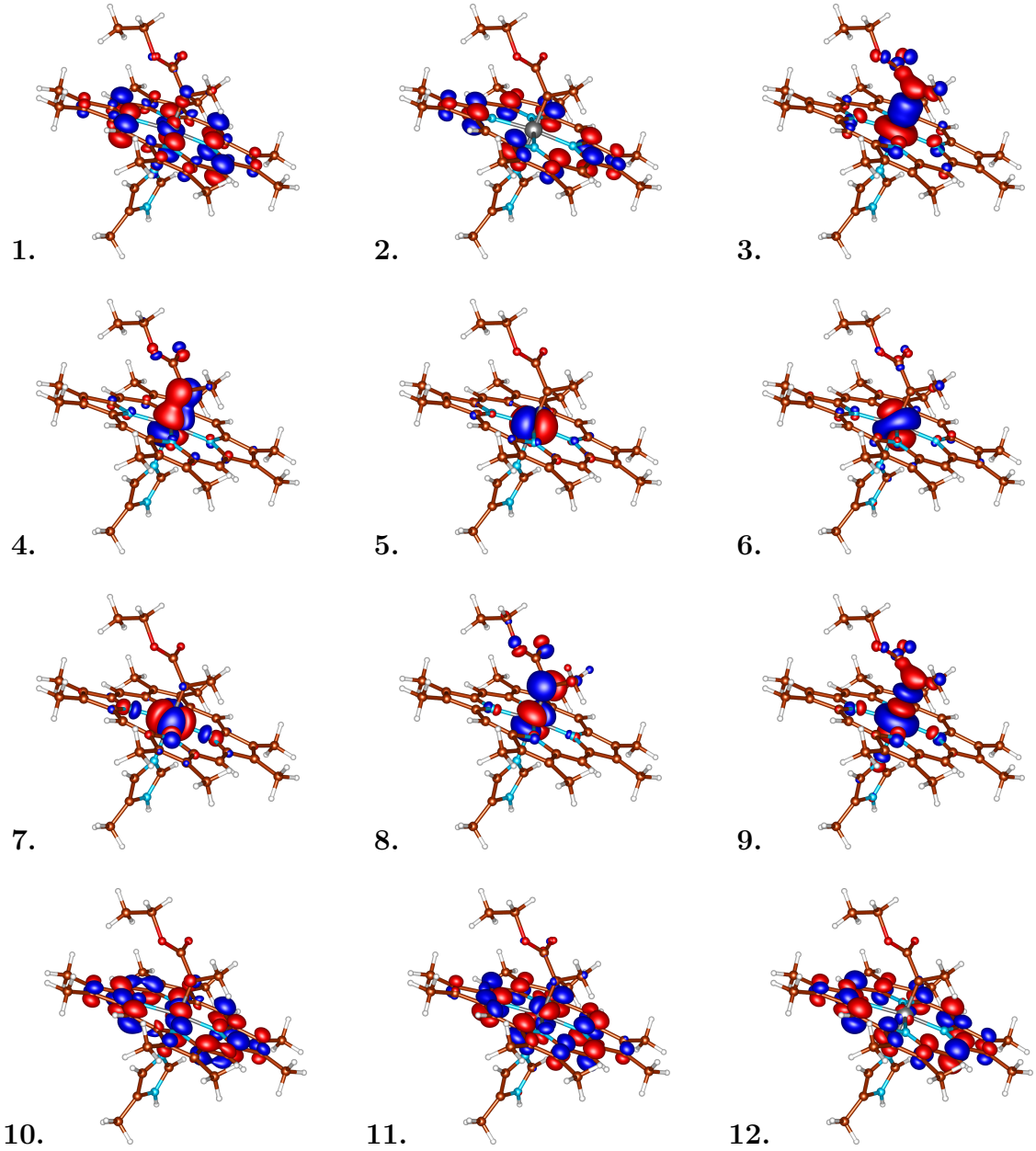

Figure 3.10: The natural orbitals from the SA-CASSCF calculations on one of the geometries. The orbitals are ordered by their natural occupation numbers.

the singlet, making it the ground state.

The comparison between the AC0 and NEVPT2 results is shown in Figure 3.11. We see that the AC0 corresponds very well with the NEVPT2 results, with the differences being less significant than the errors caused by the different basis sets. The difference is greater for the excited states, which is consistent with what is known about the behavior of the AC0 method. The method is known to be quite sensitive to proper convergence of the electronic densities. This usually leads to more accuracy if the state specific approach is used, which we were unable to do due to convergence issues.

The torsion has a significant effect, causing noticable differences even though the torsion bond is not directly included into the active space, in the sense that

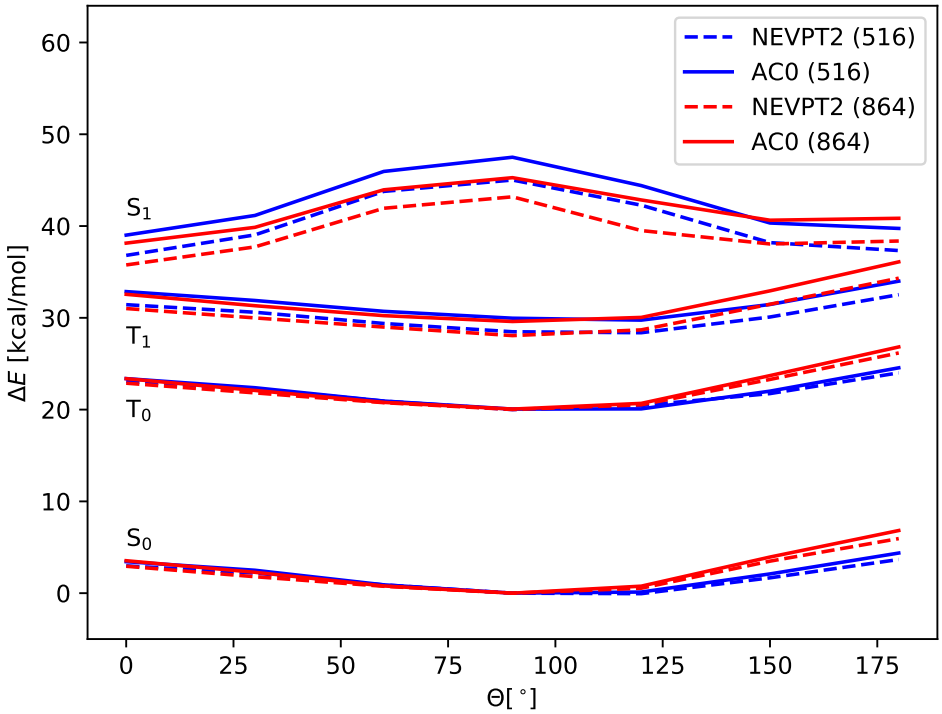

Figure 3.11: The comparison of the results of NEVPT2 and AC0, the energies for all states were shifted so that the energy of the ground state at the middle of the curve is equal to zero. Θ represents the torsion angle.

the torsion bond is not important for any of the orbitals in the active space. This well demonstrates the fine tuning ability of the porphyrin core, which is one of the most important reasons for the biochemical versatility of porphyrin systems, which consequently led to its widespread occurrence in most living organisms. This can be also heavily exploited in the engineered cytochrome transpherases, as mutations can affect both the possible reactions, and both the reaction yields and the stereoselectivity of the products.

But unfortunately, this means great complications for obtaining both accurate and practically applicable quantum chemical results on porphyrinic systems, as this will call for the inclusion of a large portion of the porphyrin surroundings, which consequently increases the computational cost of the calculations. Also, as has been mentioned earlier, these systems call for use of multireference methods, which usually exhibit quite unfavorable scaling with the number of basis set functions.

Also, although the results for this system seem pretty good, especially because the correspondence of the AC0 method with NEVPT2 is great, the results on the porphyrin model show us it will be necessary to do the CASSCF calculations with a larger active space, which will have to be done using the DMRG-CASSCF method. For the active space size necessary NEVPT2 calculations would be a problem, as cost of the method is scaling quite badly with the active space size, but AC0 should be feasible. We also hope that enlarging the active space would ensure convergence of all 4 of the states using the SS-approach, which would lead

to increased accuracy of the AC0 results. This was the case for the porphyrin model, where the CAS(12,16) CASSCF calculations struggled with serious convergence issues, which were not present for the CAS(32,34) calculations.

As can be seen from the previous paragraph, these calculations will be the subject of further research, hopefully obtaining results using a significantly larger active space, driven by our progressing work on the software implementations of the DMRG, DMRG-CASSCF and AC0 methods, which should give us further insight into this porphyrin system while simultaneously serving as a good playground for testing the capabilities of these methods.

Conclusions

On a model of the porphyrin molecule, using three different geometries and two basis sets, we have performed state specific CASSCF or DMRG-CASSCF calculations for the lowest quintet and triplet states, using three different active spaces, with the largest CAS(32,34) containing all orbitals suggested by the often stated rules regarding the active space selection. All the results were then used for subsequent correlation treatment via the recently developed TCCSD and TCCSD(T) methods.

For a larger molecule, representing an iron porphyrin carbene reaction intermediate, we have performed SA-CASSCF calculations over 4 states along one torsion coordinate, corrected with subsequent state of the art AC0 correlation treatment.

We have well illustrated the struggles with obtaining accurate results for application relevant porphyrin systems. We've shown that to obtain the correct energetical spectrum it is crucial to include the full porphyrin π-system into the active space, with also accounting for the the beyond-CAS dynamic correlation. We've demonstrated that porphyrin systems are extremely sensitive to geometrical changes, including those outside the porphyrin core.

Despite that, by using the model of the porphyrin molecule and observing the effects of geometrical changes on the energies of the lowest spin states, after accounting for both the static and dynamic correlation, using new experiments on the Fe(II)-phtalocyanine molecule, we were able to both explain the experimental results on Fe(II)-tetraphenylporphyrin and predict a quintet ground state for the isolated porphyrin molecule.

We have also found that the AC0 method is a viable option for the treatment of dynamical correlation, with a favorable cost/accuracy tradeoff, especially when large active spaces, often encountered when using the DMRG method.

While some of the obtained results are already in the process of being published [50], we have many plans for the future. In the near future we plan to perform DMRG-CASSCF calculations on the full Fe(II) porphyrin, after implementing the DMRG-CASSCF interface with the massively parallel MOLMPS program [46], recently developed in our group, which offers a great speedup compared to the previously used Budapest DMRG code. Motivated by the promising results regarding AC0 accuracy, we mean to further develop the interface with the AC0 program, allowing us to use another form of dynamical correlation treatment on top of the DMRG-CASSCF procedure. Also, as currently the interface implemented in the ORCA package is the bottleneck limiting the number of basis functions in the AC0 procedure which we are able to use, we plan to further work on improving the efficiency of the implementation.

Bibliography

[1] Michael R. Moore and Peter B. Disler. Chemistry and biochemistry of the porphyrins and porphyrias. *Clinics in Dermatology*, 3(2):7 – 23, 1985.

[2] Jun-hong Chou, Margaret E. Kosal, Hari Singh Nalwa, Neal A. Rakow, and Kenneth S. Suslick. Applications of porphyrins and metalloporphyrins to materials chemistry. 2000.

[3] Jiayuan Kou, Dou Dou, and Liming Yang. Porphyrin photosensitizers in photodynamic therapy and its applications. *Oncotarget*, 8, 07 2015.

[4] John A McIntosh, Christopher C Farwell, and Frances H Arnold. Expanding P450 catalytic reaction space through evolution and engineering. *Current Opinion in Chemical Biology*, 19:126 – 134, 2014. Biocatalysis and biotransformation Bioinorganic chemistry.

[5] Daniel Mansuy, Marc Lange, J. C. Chottard, J. F. Bartoli, Bernard Chevrier, and Raymond Weiss. Dichlorocarbene complexes of iron(ii)-porphyrins–crystal and molecular structure of fe(tpp)(ccl2)(h2o). *Angewandte Chemie International Edition in English*, 17(10):781–782, 1978.

[6] Minbo Lan, Hongli Zhao, Huihui Yuan, Chengrui Jiang, Shaohua Zuo, and Ying Jiang. Absorption and epr spectra of some porphyrins and metalloporphyrins. *Dyes and Pigments*, 74(2):357 – 362, 2007.

[7] Kevin M. Smith, Dane A. Goff, Raymond J. Abraham, and James E. Plant. The nmr spectra of porphyrins: 22—ring current effects in chlorins versus porphyrins. *Organic Magnetic Resonance*, 21(8):505–511, 1983.

[8] Blanka Vlčková, Petr Šmejkal, Martin Michl, Marek Procházka, Peter Mojzeš, František Lednický, and Jiří Pfleger. Surface-enhanced resonance raman spectroscopy of porphyrin and metalloporphyrin species in systems with ag nanoparticles and their assemblies. *Journal of Inorganic Biochemistry*, 79(1):295 – 300, 2000.

[9] Giovanni Li Manni and Ali Alavi. Understanding the mechanism stabilizing intermediate spin states in Fe(II)-porphyrin. *The Journal of Physical Chemistry A*, 122(22):4935–4947, 2018.

[10] Jan Almlof, Thomas H. Fischer, Paul G. Gassman, Abhik Ghosh, and Marco Haeser. Electron correlation in tetrapyrroles: ab initio calculations on porphyrin and the tautomers of chlorin. *The Journal of Physical Chemistry*, 97(42):10964–10970, 1993.

[11] Diane C. Rawlings, Martin Gouterman, Ernest R. Davidson, and David Feller. Theoretical investigations of the electronic states of porphyrins. iii. low-lying electronic states of porphinatoiron(ii). *International Journal of Quantum Chemistry*, 28(6):773–796, 1985.

[12] Kasper Planeta Kepp. Heme: From quantum spin crossover to oxygen manager of life. *Coordination Chemistry Reviews*, 344:363 – 374, 2017.

[13] Abhik Ghosh and Peter R. Taylor. Iron(iv) porphyrin difluoride does not exist: Implications for dft calculations on heme protein reaction pathways. *Journal of Chemical Theory and Computation*, 1(4):597–600, 2005. PMID: 26641680.

[14] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, Inc., Mineola, first edition, 1996.

[15] Steven R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356, Oct 1993.

[16] Ulrich Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.*, 326(1):96 – 192, 2011.

[17] Roger Penrose. Applications of negative-dimensional tensors. 1971.

[18] Wikipedia contributors. Penrose graphical notation — Wikipedia, the free encyclopedia, 2019. [Online; accessed 19-May-2020].

[19] Ö. Legeza, J. Röder, and B. A. Hess. Controlling the accuracy of the density-matrix renormalization-group method: The dynamical block state selection approach. *Physical Review B*, 67:125114, Mar 2003.

[20] Ernest R. Davidson. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *Journal of Computational Physics*, 17(1):87 – 94, 1975.

[21] Steven R. White and Richard L. Martin. Ab initio quantum chemistry using the density matrix renormalization group. *The Journal of Chemical Physics*, 110(9):4127–4130, 1999.

[22] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G. Kin-Lic Chan. The ab-initio density matrix renormalization group in practice. *The Journal of Chemical Physics*, 142:034102, 2015.

[23] Christopher J. Stein and Markus Reiher. Automated selection of active orbital spaces. *Journal of Chemical Theory and Computation*, 12(4):1760–1771, 2016. PMID: 26959891.

[24] G. Barcza, Ö. Legeza, K. H. Marti, and M. Reiher. Quantum-information analysis of electronic states of different molecular structures. *Phys. Rev. A*, 83:012508, Jan 2011.

[25] Valera Veryazov, Per Åke Malmqvist, and Björn O. Roos. How to select active space for multiconfigurational quantum chemistry? *International Journal of Quantum Chemistry*, 111(13):3329–3338, 2011.

[26] I. Shavitt and R.J. Bartlett. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*. Cambridge Molecular Science. Cambridge University Press, 2009.

[27] Dmitry I. Lyakh, Monika Musiał, Victor F. Lotrich, and Rodney J. Bartlett. Multireference nature of chemistry: The coupled-cluster view. *Chemical Reviews*, 112(1):182–243, 2012. PMID: 22220988.

[28] Tomoko Kinoshita, Osamu Hino, and Rodney J. Bartlett. Coupled-cluster method tailored by configuration interaction. *The Journal of Chemical Physics*, 123(7):074106, 2005.

[29] Libor Veis, Andrej Antalík, Jiří Brabec, Frank Neese, Örs Legeza, and Jiří Pittner. Coupled cluster method with single and double excitations tailored by matrix product state wave functions. *The Journal of Physical Chemistry Letters*, 7(20):4072–4078, October 2016.

[30] Frank Neese, Andreas Hansen, and Dimitrios G. Liakos. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *The Journal of Chemical Physics*, 131(6):064103, 2009.

[31] Christoph Riplinger and Frank Neese. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *The Journal of Chemical Physics*, 138(3):034106, 2013.

[32] Frank Neese, Frank Wennmohs, and Andreas Hansen. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *The Journal of Chemical Physics*, 130(11):114108, 2009.

[33] Jakub Lang, Andrej Antalík, Libor Veis, Jan Brandejs, Jiří Brabec, Örs Legeza, and Jiří Pittner. Near-linear scaling in DMRG-based tailored coupled clusters: An implementation of DLPNO-TCCSD and dlpno-tccsd(t).

[34] J. Harris. Adiabatic-connection approach to kohn-sham theory. *Phys. Rev. A*, 29:1648–1659, Apr 1984.

[35] Katarzyna Pernal. Electron correlation from the adiabatic connection for multireference wave functions. *Phys. Rev. Lett.*, 120:013001, Jan 2018.

[36] Ewa Pastorczak and Katarzyna Pernal. Correlation energy from the adiabatic connection formalism for complete active space wave functions. *Journal of Chemical Theory and Computation*, 14(7):3493–3503, 2018. PMID: 29787257.

[37] Ewa Pastorczak and Katarzyna Pernal. Electronic excited states from the adiabatic-connection formalism with complete active space wave functions. *The Journal of Physical Chemistry Letters*, 9(18):5534–5538, 2018. PMID: 30192553.

[38] Dana Nachtigallová, Andrej Antalík, Rabindranath Lo, Robert Sedlák, Debashree Manna, Jiří Tuček, Juri Ugolotti, Libor Veis, Örs Legeza, Jiří Pittner, Radek Zbořil, and Pavel Hobza. An isolated molecule of iron(ii) phthalocyanin exhibits quintet ground-state: A nexus between theory and experiment. *Chemistry – A European Journal*, 24(51):13413–13417, 2018.

[39] G. Lang, K. Spartalian, Christopher A. Reed, and James P. Collman. Mössbauer effect study of the magnetic properties of s=1 ferrous tetraphenylporphyrin. *The Journal of Chemical Physics*, 69(12):5424–5427, 1978.

[40] Kristine Pierloot, Quan Manh Phung, and Alex Domingo. Spin state energetics in first-row transition metal complexes: contribution of (3s3p) correlation and its description by second-order perturbation theory. *Journal of chemical theory and computation*, 13(2):537–553, 2017.

[41] Martin Gouterman, Georges H. Wagnière, and Lawrence C. Snyder. Spectra of porphyrins: Part II. four orbital model. *Journal of Molecular Spectroscopy*, 11(1):108 – 127, 1963.

[42] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.*, 7:3297–3305, 2005.

[43] M. J. Frisch et al. Gaussian˜16 Revision C.01, 2016. Gaussian Inc. Wallingford CT.

[44] Frank Neese. The ORCA program system. *WIREs Computational Molecular Science*, 2(1):73–78, 2012.

[45] Ö. Legeza, J. Röder, and B. A. Hess. Controlling the accuracy of the density-matrix renormalization-group method: The dynamical block state selection approach. *Phys. Rev. B*, 67:125114, Mar 2003.

[46] Jiri Brabec, Jan Brandejs, Karol Kowalski, Sotiris Xantheas, Ors Legeza, and Libor Veis. Massively parallel quantum chemical density matrix renormalization group method.

[47] Russell D. Lewis, Marc Garcia-Borràs, Matthew J. Chalkley, Andrew R. Buller, K. N. Houk, S. B. Jennifer Kan, and Frances H. Arnold. Catalytic iron-carbene intermediate revealed in a cytochrome c carbene transferase. *Proceedings of the National Academy of Sciences*, 115(28):7308–7313, 2018.

[48] Martin Gouterman and Gamal Khalil. Porphyrin free base phosphorescence. *Journal of Molecular Spectroscopy*, 53:88–100, 10 1974.

[49] Rashid Valiev, V. Cherepanov, Gleb Baryshnikov, and Dage Sundholm. First-principles method for calculating the rate constants of internal-conversion and intersystem-crossing transitions. *Physical Chemistry Chemical Physics*, 20, 02 2018.

[50] Andrej Antalík, Dana Nachtigallová, Rabindranath Lo, Mikuláš Matoušek, Jakub Lang, Örs Legeza, Jiří Pittner, Pavel Hobza, and Libor Veis. Ground state of the Fe(II)-porphyrin model system corresponds to the quintet state: A DFT and DMRG-based tailored CC study, 2020.

List of Figures

A. Attachments

A.1 The used geometries of the porphyrine model

A.1.1 Reference geometry

Available from the supplementary information of [9]

Fe	0.000000	0.000000	0.000000
N	1.406727	1.406727	0.000000
N	-1.406727	1.406727	0.000000
N	1.406727	-1.406727	0.000000
N	-1.406727	-1.406727	0.000000
C	-0.000000	3.400142	0.000000
C	-0.000000	-3.400142	0.000000
C	3.400142	-0.000000	0.000000
C	-3.400142	-0.000000	0.000000
C	1.222770	2.760387	0.000000
C	-1.222770	2.760387	0.000000
C	1.222770	-2.760387	0.000000
C	-1.222770	-2.760387	0.000000
C	2.760387	1.222770	0.000000
C	-2.760387	1.222770	0.000000
C	2.760387	-1.222770	0.000000
C	-2.760387	-1.222770	0.000000
H	0.000000	4.482672	0.000000
H	0.000000	-4.482672	0.000000
H	4.482672	0.000000	0.000000
H	-4.482672	0.000000	0.000000
H	2.181081	3.277651	0.000000
H	-2.181081	3.277651	0.000000
H	2.181081	-3.277651	0.000000
H	-2.181081	-3.277651	0.000000
H	3.277651	2.181081	0.000000
H	-3.277651	2.181081	0.000000
H	3.277651	-2.181081	0.000000
H	-3.277651	-2.181081	0.000000

A.1.2 Triplet optimized

Fe	-0.000000000	-0.000000000	0.000000000
N	1.448270867	1.448270867	0.000000000
N	-1.448270867	1.448270867	0.000000000
N	1.448270867	-1.448270867	0.000000000
N	-1.448270867	-1.448270867	0.000000000
C	-0.000000000	3.428797240	0.000000000
C	0.000000000	-3.428797240	0.000000000
C	3.428797240	-0.000000000	0.000000000
C	-3.428797240	0.000000000	0.000000000
C	1.221237955	2.795853304	0.000000000
C	-1.221237955	2.795853304	0.000000000
C	1.221237955	-2.795853304	0.000000000
C	-1.221237955	-2.795853304	0.000000000
C	2.795853304	1.221237955	0.000000000
C	-2.795853304	1.221237955	0.000000000
C	2.795853304	-1.221237955	0.000000000
C	-2.795853304	-1.221237955	0.000000000
H	0.000000000	4.514607676	-0.000000000
H	0.000000000	-4.514607676	-0.000000000
H	4.514607676	-0.000000000	-0.000000000
H	-4.514607676	0.000000000	-0.000000000
H	2.100989893	3.440586200	-0.000000000
H	-2.100989893	3.440586200	-0.000000000
H	2.100989893	-3.440586200	-0.000000000
H	-2.100989893	-3.440586200	-0.000000000
H	3.440586200	2.100989893	-0.000000000
H	-3.440586200	2.100989893	-0.000000000
H	3.440586200	-2.100989893	-0.000000000
H	-3.440586200	-2.100989893	-0.000000000

A.1.3 Quintet optimized

Fe	-0.000003334	-0.000004483	-0.000098650
N	1.548142762	1.533769649	-0.016184178
N	-1.548640753	1.534274738	0.016173811
N	1.548637499	-1.534277783	0.016177915
N	-1.548141343	-1.533769178	-0.016178698
C	-0.000141228	3.446356507	-0.000054123
C	0.000142840	-3.446357499	-0.000055635
C	3.467596490	-0.000155678	0.000050079
C	-3.467597371	0.000157359	0.000047038
C	1.254681160	2.856378941	0.015004151
C	-1.255157041	2.856784188	-0.015173535
C	1.255159484	-2.856788013	-0.015179157
C	-1.254678834	-2.856378181	0.015009758
C	2.874717233	1.253448950	-0.013990771
C	-2.875113284	1.253884771	0.014133401
C	2.875108902	-1.253881388	0.014142394
C	-2.874715761	-1.253446251	-0.013985164
H	0.000007292	4.533512778	-0.000083340
H	-0.000008418	-4.533513863	-0.000089391
H	4.554836054	0.000010488	0.000115649
H	-4.554837053	-0.000012239	0.000110716
H	2.107723495	3.541495350	0.049148103
H	-2.108100232	3.542069871	-0.049547428
H	2.108103559	-3.542071348	-0.049565150
H	-2.107721520	-3.541493962	0.049154329
H	3.554267120	2.111248485	-0.008046938
H	-3.554851777	2.111619876	0.008496239
H	3.554848963	-2.111615730	0.008509309
H	-3.554264904	-2.111246353	-0.008040732

A.2 The electronic energies for the porphyrin model

A.2.1 def2-SVP basis set

		CASSCF	TCCSD	LPNO-TCCSD	DLPNO-TCCSD	TCCSD(T)	DLPNO-TCCSD(T)
					CAS(8,12)		
Reference geometry	-triplet	-1940.88659	-1943.46640	-1943.45026	-1943.46161	-1943.58662	-1943.57860
	-quintet	-1940.90392	-1943.47319	-1943.45671	-1943.46934	-1943.59215	-1943.58517
Triplet geometry	-triplet	-1940.93620	-1943.50546	-1943.48892	-1943.50123	-1943.62435	-1943.61712
	-quintet	-1940.96703	-1943.52548	-1943.50884	-1943.52133	-1943.64313	-1943.63602
Quintet geometry	-triplet	-1940.94324	-1943.50517	-1943.49015	-1943.50152	-1943.62420	-1943.61780
	-quintet	-1940.99655	-1943.54780	-1943.53439	-1943.54404	-1943.66607	-1943.65960
					CAS(12,16)		
Reference geometry	-triplet	-1940.88659	-1943.46640	-1943.45026	-1943.46161	-1943.58662	-1943.57860
	-quintet	-1940.90392	-1943.47319	-1943.45671	-1943.46934	-1943.59215	-1943.58517
Triplet geometry	-triplet	-1940.93620	-1943.50546	-1943.48892	-1943.50123	-1943.62435	-1943.61712
	-quintet	-1940.96703	-1943.52548	-1943.50884	-1943.52133	-1943.64313	-1943.63602
Quintet geometry	-triplet	-1940.94324	-1943.50517	-1943.49015	-1943.50152	-1943.62420	-1943.61780
	-quintet	-1940.99655	-1943.54780	-1943.53439	-1943.54404	-1943.66607	-1943.65960

		DMRG-CASSCF	DMRG-DBSS	TCCSD	LPNO-TCCSD	DLPNO-TCCSD	TCCSD(T)	DLPNO-TCCSD(T)
					CAS(32,34)			
Reference geometry	-triplet	-1941.20789	-1941.21955	-1943.53927	-1943.52996	-1943.53405	-1943.62084	-1943.61219
	-quintet	-1941.20867	-1941.21973	-1943.54330	-1943.53373	-1943.53793	-1943.62424	-1943.61550
Triplet geometry	-triplet	-1941.26442	-1941.26782	-1943.57832	-1943.56941	-1943.57343	-1943.65851	-1943.65037
	-quintet	-1941.27271	-1941.28041	-1943.59393	-1943.58479	-1943.58873	-1943.67370	-1943.66516
Quintet geometry	-triplet	-1941.26949	-1941.27493	-1943.57613	-1943.56767	-1943.57134	-1943.65559	-1943.64771
	-quintet	-1941.30357	-1941.31291	-1943.61601	-1943.60732	-1943.61118	-1943.69509	-1943.68712

A.2.2 def2-TZVP basis set

		CASSCF	LPNO-TCCSD	DLPNO-TCCSD	DLPNO-TCCSD(T)	
				CAS(8,12)		
Reference geometry	-triplet	-1941.80401	-1944.91706	-1944.92891	-1945.09147	
	-quintet	-1941.82126	-1944.92124	-1944.93331	-1945.09450	
Triplet geometry	-triplet	-1941.85242	-1944.95264	-1944.96320	-1945.12557	
	-quintet	-1941.88291	-1944.97004	-1944.98180	-1945.14265	
Quintet geometry	-triplet	-1941.85626	-1944.94641	-1944.95734	-1945.12010	
	-quintet	-1941.90878	-1944.98679	-1944.99730	-1945.15731	
				CAS(12,14)		
Reference geometry	-triplet	-1941.84602	-1944.93550	-1944.94585	-1945.10279	
	-quintet	-1941.86296	-1944.94002	-1944.94987	-1945.10521	
Triplet geometry	-triplet	-1941.89414	-1944.96635	-1944.98085	-1945.13589	
	-quintet	-1941.92440	-1944.98787	-1944.99796	-1945.15131	
Quintet geometry	-triplet	-1941.89680	-1944.96239	-1944.97195	-1945.12629	
	-quintet	-1941.94909	-1945.00061	-1945.01214	-1945.16473	

		DMRG-CASSCF	DMRG-DBSS	LPNO-TCCSD	DLPNO-TCCSD	DLPNO-TCCSD(T)
				CAS(32,34)		
Reference geometry	-triplet	-1942.12238	-1942.13227	-1945.00274	-1945.00800	-1945.12911
	-quintet	-1942.11955	-1942.12829	-1945.00306	-1945.00858	-1945.12894
Triplet geometry	-triplet	-1942.17481	-1942.17854	-1945.03699	-1945.04236	-1945.16213
	-quintet	-1942.18040	-1942.18861	-1945.05061	-1945.05570	-1945.17462
Quintet geometry	-triplet	-1942.17969	-1942.18288	-1945.02826	-1945.03282	-1945.15163
	-quintet	-1942.21037	-1942.21601	-1945.06416	-1945.06938	-1945.18753

A.3 The lowest excited states of the IPC

The dominant contributions to the lowest SA-CASSCF states for the iron porphyrin carbene system. The orbital occupancies are refering to the orbitals in figure 3.10 in the same order.

Triplet states		
0		E= -3164.539
	222212100000	0.539
	222112110000	0.102
1		E= -3164.525
	222221100000	0.533
	222121110000	0.110
2		E= -3164.4944
	222211110000	0.639
	222111120000	0.167

Singlet states		
0		E= -3164.5377
	222222000000	0.732
	222022020000	0.086
1		E= -3164.4839
	222221010000	0.319
	222212100000	0.283
	222121020000	0.075
2		E= -3164.4746
	222212010000	0.431
	222221100000	0.192
3		E= -3164.4677
	222212100000	0.374
	222221010000	0.266
	222121020000	0.060
4		E= -3164.4457
	212212100100	0.460
	212112110100	0.099
5		E= -3164.4326
	212221100010	0.349
	212221100100	0.149
6		E= -3164.4296
	212212100010	0.441
	212212100100	0.097
7		E= -3164.4294
	212221100100	0.384
	212221100010	0.114
8		E= -3164.4234
	222221100000	0.423
	222212010000	0.126
9		E= -3164.4170
	222211110000	0.526
	222111120000	0.121
10		E= -3164.4166
	222122100000	0.330
	222202110000	0.150
11		E= -3164.4155
	122212100100	0.414
	122112110100	0.087
12		E= -3164.4060
	212222000100	0.414
	122222000010	0.19052
13		E= -3164.4040
	222211110000	0.410
	222111120000	0.094
14		E= -3164.4035
	122221100010	0.310
	122221100100	0.135
15		E= -3164.4027
	212222000010	0.341
	122222000100	0.246
	212222000100	0.051
16		E= -3164.4010
	222202110000	0.167
	212211110010	0.110
	222211110000	0.103
	222220110000	0.069
17		E= -3164.4009
	212211110010	0.273
	212211110100	0.091
18		E= -3164.4005
	122221100100	0.353
	122221100010	0.151
19		E= -3164.3998
	212122010100	0.254
	212221010010	0.199
	212222010000	0.196
	212122010010	0.041

A.4 Comparison of the basis sets for the IPC

The SA-CASSCF energies (in miliHartrees) for the 4 lowest spin states using the given basis set. The number in parenthesis under the basis set denote the number of basis set functions. In bold are the two basis sets used in our calculations.

	cc-pVQZ (3839)	cc-pVTZ (2006)	cc-pVDZ (883)	**Larger (816)**	6-31G(d) (748)	**Smaller (564)**	6-31G (516)
S_0	0	0	0.081	**1.3**	0.54	**0.9**	3.7
T_0	0.1	0.013	0	**0**	0	**0**	0
T_1	15.5	15.5	14.9	**14.4**	14.1	**15.4**	15.1
S_1	55.9	55.9	55.7	**55.1**	53.7	**56.2**	53.3